

**UNIVERSIDADE FEDERAL DO ABC**  
**CENTRO DE CIÊNCIAS NATURAIS E HUMANAS**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOCIÊNCIA**

**SAMUEL PERES CHAGAS**

**SUPEREXPRESSÃO DE GENES ENVOLVIDOS NA BIOSÍNTESE DE ÁCIDO  
ROSMARÍNICO EM SETARIA VIRIDIS PARA REDUÇÃO DA RECALCITRÂNCIA  
DA BIOMASSA LIGNOCELULÓSICA**

São Bernardo do Campo – SP

2025

SAMUEL PERES CHAGAS

**SUPEREXPRESSÃO DE GENES ENVOLVIDOS NA BIOSÍNTESE DE ÁCIDO  
ROSMARÍNICO EM SETARIA VIRIDIS PARA REDUÇÃO DA RECALCITRÂNCIA  
DA BIOMASSA LIGNOCELULÓSICA**

Tese apresentada ao Programa de Pós-graduação em Biotecnologia da Universidade Federal do ABC como requisito à obtenção do título de Doutor em Biotecnologia. Linha de pesquisa: Propriedades físicas, químicas e biológicas de moléculas complexas.

Orientador: Prof. Dr. Wagner Rodrigo de Souza

Co-orientador: Dr. José Franciraldo de Lima

São Bernardo do Campo – SP

2025

## FICHA CATALOGRÁFICA

**Sistema de Bibliotecas da Universidade Federal do ABC**  
Elaborada pelo Sistema de Geração de Ficha Catalográfica da UFABC  
com os dados fornecidos pelo(a) autor(a).

Samuel, Peres Chagas

Superexpressão de genes envolvidos na biossíntese de Ácido Rosmarínico em *Setaria viridis* para redução da recalcitrância da biomassa lignocelulósica / Peres Chagas Samuel. — 2025.

121 fls.

Orientação de: Wagner Rodrigo de Souza

Coorientação de: José Franciraldo de Lima

Tese (Doutorado) — Universidade Federal do ABC, Programa de Pós-Graduação em Biotecnociência, São Bernardo do Campo, 2025.

1. Gramíneas. 2. Parede celular. 3. Metabolismo fenólico. 4. Lignina. 5. Etanol de segunda geração. I. Souza, Wagner Rodrigo de. II. Lima, José Franciraldo de. III. Programa de Pós-Graduação em Biotecnociência, 2025. IV. Título.

## DECLARAÇÃO

Este exemplar foi alterado e revisado em relação à versão original, de acordo com as observações levantadas pela banca no dia da defesa, sob responsabilidade única do autor e com a anuência do(s) seu(s) orientador(es).

**Santo André, 26 de Abril de 2025.**

*Samuel P. Chagas*

---

Samuel Peres Chagas

Autor

Documento assinado digitalmente  
 **WAGNER RODRIGO DE SOUZA**  
Data: 26/03/2025 11:57:06-0300  
Verifique em <https://validar.itl.gov.br>

---

Prof. Dr. Wagner Rodrigo de Souza

Orientador

Documento assinado digitalmente  
 **JOSE FRANCIRALDO DE LIMA**  
Data: 26/03/2025 10:33:00-0300  
Verifique em <https://validar.itl.gov.br>

---

Prof. Dr. José Franciraldo de Lima

Co-orientador



**MINISTÉRIO DA EDUCAÇÃO**

**Fundação Universidade Federal do ABC**

Avenida dos Estados, 5001 - Bairro Santa Terezinha - Santo André - SP  
CEP 09210-580 · Fone: (11) 4996-0017

**Ata de Defesa de Tese de Doutorado e Folha de Assinaturas**

No dia 18 de Março de 2025 às 14:00, no local: <https://conferenciaweb.rnp.br/sala/wagner-50>, realizou-se a Defesa da Tese de Doutorado, que constou da apresentação do trabalho intitulado " SUPEREXPRESSÃO DE GENES ENVOLVIDOS NA BIOSÍNTESE DE ÁCIDO ROSMARÍNICO EM SETARIA VIRIDIS PARA REDUÇÃO DA RECALCITRÂNCIA DA BIOMASSA LIGNOCELULÓSICA" de autoria do candidato, SAMUEL PERES CHAGAS, RA nº 23202010900, discente do Programa de Pós-Graduação em BIOTECNOCIÊNCIA da UFABC, sob orientação do Profº WAGNER RODRIGO DE SOUZA. Concluídos os trabalhos de apresentação e arguição, o candidato foi considerado APROVADO pela Banca Examinadora.

E, para constar, foi lavrada a presente ata e folha de assinaturas assinada pelos membros da Banca.

Documento assinado digitalmente  
 **WAGNER RODRIGO DE SOUZA**  
Data: 21/03/2025 15:37:31-0300  
Verifique em <https://validar.iti.gov.br>

---

**Dr. WAGNER RODRIGO DE SOUZA, UFABC**  
Presidente - Interno ao Programa

Documento assinado digitalmente  
 **NATHALIA DE SETTA COSTA**  
Data: 21/03/2025 14:46:46-0300  
Verifique em <https://validar.iti.gov.br>

---

**Dra. NATHALIA DE SETTA COSTA, UFABC**  
Membro Titular - Examinador(a) Interno ao Programa

Documento assinado digitalmente  
 **HANA PAULA MASUDA**  
Data: 20/03/2025 18:47:35-0300  
Verifique em <https://validar.iti.gov.br>

---

**Dra. HANA PAULA MASUDA, UFABC**  
Membro Titular - Examinador(a) Externo ao Programa

Documento assinado digitalmente  
 **ELISSON ANTONIO DA COSTA ROMANEL**  
Data: 20/03/2025 18:07:04-0300  
Verifique em <https://validar.iti.gov.br>

---

**Dr. ELISSON ROMANEL, USP**  
Membro Titular - Examinador(a) Externo à Instituição



**MINISTÉRIO DA EDUCAÇÃO**

**Fundação Universidade Federal do ABC**

Avenida dos Estados, 5001 - Bairro Santa Terezinha - Santo André - SP  
CEP 09210-580 · Fone: (11) 4996-0017

Documento assinado digitalmente



**WANDERLEY DANTAS DOS SANTOS**

Data: 20/03/2025 19:08:57-0300

Verifique em <https://validar.iti.gov.br>

---

**Dr. WANDERLEY DANTAS DOS SANTOS, UEM**  
Membro Titular - Examinador(a) Externo à Instituição

---

**Dra. ELIZABETH TEODOROV, UFABC**  
Membro Suplente - Examinador(a) Interno ao Programa

---

**Dr. KAROLINE ESTEFANI DUARTE**  
Membro Suplente - Examinador(a) Externo à Instituição

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001

## **AGRADECIMENTOS**

Este trabalho é resultado da colaboração de inúmeras pessoas que tornaram possível a realização dos experimentos, que fizeram parte das discussões científicas, e também que deram suporte emocional durante os últimos 4 anos e meio. Agradeço aos meus pais e à minha família, que mesmo a distância, sempre me deram suporte e acreditaram que a caminhada seria valiosa.

Agradeço primeiramente ao meu orientador, prof. Wagner R. de Souza, pelo apoio e suporte contínuo, pelas discussões científicas e por acreditar na minha capacidade. Agradeço aos demais membros do grupo de pesquisa, Fernanda O. Menezes, Tamiris M. Kobayasi, Pollyanna Castilho, e Karoline Duarte, responsável por transmitir grande parte do conhecimento prático necessário para realização deste trabalho.

Agradeço também aos professores do laboratório de Biotecnologia Verde, de laboratórios vizinhos e Central Multiusuário de São Bernardo do Campo, dos quais fiz uso de vários equipamentos que permitiram a obtenção dos resultados, como prof. Danilo Centeno, prof. Nathalia de Setta, prof. César Augusto e prof. Sérgio Sasaki. Agradeço ainda Wout Boerjan, Dyoni Oliveira e Thatiane Mota, por todo o apoio, experiência transmitida e discussões científicas realizadas durante o período sanduíche no Instituto de Biotecnologia do Flandres.

## **AGRADECIMENTOS**

Agradeço a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, pelo financiamento do projeto de doutorado sanduíche CAPES/PRINT - Edital nº 41/2017, número de processo 88887.717223/2022-00.

Agradeço a Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP; a Financiadora de Estudos e Projetos – FINEP; e o Ministério da Ciência, Tecnologia e Inovação – MCTI; que por meio do Programa de Recursos Humanos da ANP para o Setor Petróleo e Gás – PRH-ANP/MCTI 49.1 deu todo o suporte financeiro necessário para a realização do trabalho.

O presente trabalho também foi realizado com apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Brasil. Processo nº 2019/04878-7, referente ao projeto do prof. Wagner Rodrigo de Souza.

Esta pesquisa utilizou instalações do Laboratório Nacional de Biorrenováveis (LNBR), do Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), uma Organização Social supervisionada pelo Ministério da Ciência, Tecnologia e Inovação (MCTI). A equipe da instalação aberta de Multi-ômicas (OMICS) é reconhecida pela assistência durante os experimentos (20250236).

Agradeço, por fim, a Universidade Federal do ABC, ao centro de Ciências Naturais e Humanas (CCNH) e ao programa de pós-graduação em Biotecnociência.

## RESUMO

A crescente demanda energética e o compromisso com a redução das emissões atmosféricas faz com que novas alternativas sejam estudadas para suprir as necessidades energéticas da nossa sociedade atual. Nesse cenário, o etanol de segunda geração (E2G) surge como uma alternativa renovável e de baixo impacto ambiental, principalmente por ter como principal matéria prima a biomassa lignocelulósica, que é bastante abundante sob a forma de resíduos agroindustriais. Porém, a produção de etanol 2G encontra alguns gargalos, principalmente voltados ao custo de produção, que precisa de duas etapas adicionais se comparado ao etanol 1G: pré-tratamento da biomassa e hidrólise enzimática. Isto tem reflexos no custo do produto final, que ainda precisa de desenvolvimentos para se tornar competitivo e mais atrativo ao mercado. Neste aspecto, surgem alternativas para a redução dos custos, e uma delas é a modificação genética da biomassa lignocelulósica, mais precisamente da lignina, presente na parede celular das plantas, a fim de tornar a biomassa mais suscetível aos pré-tratamentos e reduzir o custo do processo. Uma alternativa é a modificação da composição monomérica da lignina, através da incorporação de compostos como o ácido rosmarínico (AR), que pode atuar como um monômero mais solúvel, tornando o processo de pré-tratamento mais eficiente. Sendo assim, este trabalho se propõe a estudar a superexpressão dos genes tirosina aminotransferase de *Salvia miltiorrhiza* (*SmTAT*) e hidroxifenilpiruvato redutase de *Setaria viridis* (*SvHPPR*), envolvidos na biossíntese de AR em *Setaria viridis*, uma planta modelo para gramíneas. Nossa hipótese é que a superprodução do AR leve à sua incorporação na lignina, introduzindo ligações éster no polímero que não comprometam o papel biológico da lignina, mas que possam ser facilmente quebradas durante o pré-tratamento, facilitando a remoção da lignina e promovendo a sacarificação. Os genes foram inseridos nas plantas através de transformação genética baseada em *Agrobacterium tumefaciens*, confirmadas via PCR e avaliadas quanto à expressão gênica por qPCR. A caracterização da biomassa das plantas transgênicas indicou maior eficiência de sacarificação, chegando a valores até 70% superiores se comparados às plantas selvagens em linhagens com maior presença de AR, além de quantidades de lignina alteradas, assim como sua composição. Tais fatores reforçam a hipótese de que alterações no metabolismo fenólico de gramíneas podem contribuir para o desenvolvimento de cultivares mais suscetíveis a pré-tratamentos, para melhorar a eficiência da produção do etanol 2G.

**Palavras-chave:** Gramíneas, Parede Celular, Metabolismo Fenólico, Lignina, Etanol de Segunda Geração.

## ABSTRACT

The growing energy demand and the commitment to reducing atmospheric emissions have led to the exploration of new alternatives to meet the energy needs of our modern society. In this scenario, second-generation ethanol (2G ethanol) emerges as a renewable and low-environmental-impact alternative, mainly due to its primary raw material—lignocellulosic biomass—which is abundantly available in the form of agro-industrial residues. However, 2G ethanol production faces several bottlenecks, particularly regarding production costs, as it requires two additional steps compared to first-generation ethanol (1G ethanol): biomass pretreatment and enzymatic hydrolysis. These additional steps impact the final product cost, which still requires further developments to become more competitive and attractive to the market. In this regard, alternatives for cost reduction have been explored, one of which is the genetic modification of lignocellulosic biomass, specifically lignin, which is present in the plant cell wall. The goal is to make the biomass more susceptible to pretreatments and reduce processing costs. One approach is to modify the monomeric composition of lignin by incorporating compounds such as rosmarinic acid (RA), which can act as a more soluble monomer, making the pretreatment process more efficient. Therefore, this study aims to investigate the overexpression of the genes tyrosine aminotransferase from *Salvia miltiorrhiza* (*SmTAT*) and hydroxyphenylpyruvate reductase from *Setaria viridis* (*SvHPPR*), both involved in the biosynthesis of RA in *Setaria viridis*, a model plant for grasses. Our hypothesis is that RA overproduction leads to its incorporation into lignin, introducing ester bonds in the polymer that do not compromise lignin's biological role but can be easily broken during pretreatment, facilitating lignin removal and enhancing saccharification. The genes were introduced into plants via *Agrobacterium tumefaciens*-mediated genetic transformation, confirmed by PCR, and analyzed for gene expression through qPCR. Biomass characterization of transgenic plants indicated higher saccharification efficiency, reaching up to 70% higher values compared to wild-type plants in lines with increased RA levels, along with altered lignin content and composition. These factors support the hypothesis that modifications in the phenolic metabolism of grasses can contribute to the development of cultivars more susceptible to pretreatment, ultimately improving the efficiency of 2G ethanol production.

**Key-words:** Grasses, Cell Wall, Phenolic Metabolism, Lignin, Second Generation Ethanol.

## Índice de Figuras

<b>Figura 1</b> – Matriz lignocelulósica da parede celular secundária de gramíneas.....	21
<b>Figura 2</b> – Estrutura da macromolécula de lignina.....	23
<b>Figura 3</b> – Modelo para a rota de biossíntese de lignina em gramíneas.....	25
<b>Figura 4</b> – Estrutura do Ácido Rosmarínico.....	29
<b>Figura 5</b> – Rota de biossíntese de Ácido Rosmarínico.....	31
<b>Figura 6</b> – Esquema de construção dos vetores de expressão.....	37
<b>Figura 7</b> – Etapas do processo de transformação genética de <i>S.viridis</i> por <i>Agrobacterium tumefaciens</i> .....	41
<b>Figura 8</b> – Etapas do processo de transformação genética de <i>S. viridis</i> .....	62
<b>Figura 9</b> – Gel de eletroforese da reação de PCR com primers pAct dos eventos transformados de E430 e E432.....	63
<b>Figura 10</b> – Expressão gênica de <i>SvHPPR</i> e <i>SmTAT</i> para folhas de linhagens E430.....	64
<b>Figura 11</b> – Expressão gênica <i>SvHPPR</i> para colmos de linhagens E432.....	65
<b>Figura 12</b> – Hidrólise enzimática da biomassa bruta de colmos de linhagens E430 em condição sem pré-tratamento.....	66
<b>Figura 13</b> – Hidrólise enzimática de CWR de colmos de linhagens E430 em condição sem pré-tratamento.....	67
<b>Figura 14</b> – Hidrólise enzimática de biomassa bruta de colmos de linhagens E432 em condição sem pré-tratamento.....	68
<b>Figura 15</b> – Quantificação de açúcares pelos métodos Benedict e Glicose Liquiform.....	69
<b>Figura 16</b> – Hidrólise enzimática de CWR de colmos de linhagens E432 em condição sem pré-tratamento.....	70
<b>Figura 17</b> – Caracterização da parede celular a partir do conteúdo de lignina, celulose e hemicelulose de linhagens de E430.....	71
<b>Figura 18</b> – Caracterização da parede celular a partir do conteúdo de lignina, celulose e hemicelulose de linhagens de E432.....	72
<b>Figura 19</b> – Composição de lignina e razão S/G de linhagens E430.....	74
<b>Figura 20</b> – Composição de lignina e teores de S/G de linhagens E432.....	75
<b>Figura 21</b> – Fenólicos totais e AR de colmos da construção E430.....	76
<b>Figura 22</b> – Fenólicos totais e AR de colmos da construção E432.....	77
<b>Figura 23</b> – Biomassa e Inflorescências de linhagens E430 e E432.....	76
<b>Figura 24</b> – Comparação da abundância absoluta dos íons de compostos das linhagens das construções em comparação com plantas selvagens em folhas e colmos.....	79

## **Lista de Tabelas**

<b>Tabela 1</b> – Enzimas chave na biossíntese de Ácido Rosmarínico.....	61
<b>Tabela 2</b> – Anotação de possíveis compostos diferencias de E430 e E432.....	80

## **Índice de Equações**

<b>Equação 1</b> - Cálculo do qui-quadrado.....	44
<b>Equação 2</b> - Cálculo da expressão gênica pelo $\Delta\Delta Ct$ .....	46
<b>Equação 3</b> - Cálculo do conteúdo de lignina método brometo de acetila.....	49
<b>Equação 4</b> - Cálculo do conteúdo de lignina método CASA.....	50

## Sumário

<b>1. Introdução Geral.....</b>	<b>16</b>
<b>2. Referencial Teórico.....</b>	<b>19</b>
2.1 Panorama do etanol de segunda geração.....	19
2.2 Composição da parede celular de plantas.....	20
2.3 Biossíntese e o papel da lignina na parede celular.....	22
2.4 Engenharia genética da rota de biossíntese de lignina.....	26
2.5 O Ácido Rosmarínico como um monômero de lignina.....	27
2.6 A presença do Ácido Rosmarínico na família Poaceae e em <i>Setaria viridis</i> .....	31
2.7 Utilização de <i>Setaria viridis</i> como modelo para cana-de-açúcar.....	32
<b>3. Justificativa e Hipóteses.....</b>	<b>33</b>
<b>4. Objetivos.....</b>	<b>34</b>
<b>5. Metodologia.....</b>	<b>35</b>
5.1 Estudo dos genes da maquinaria de AR em <i>S. viridis</i> e construção dos vetores de expressão.....	35
5.2 Transformação genética de <i>S. viridis</i> .....	39
5.3 Extração de DNA e confirmação via PCR.....	41
5.4 Avanço de geração para obtenção de plantas homozigotas.....	43
5.5 Extração e tratamento de RNA, síntese de cDNA e PCR quantitativa em tempo real.....	45
5.6 Coleta do tecido vegetal para análises bioquímicas e parâmetros biomassa e inflorescências.....	47
5.7 Preparo da parede celular (CWR).....	48
5.8 Lignina pelo método de Brometo de Acetila.....	49
5.9 Lignina pelo método CASA.....	50
5.10 Tioacidólise.....	51
5.11 (Hemi)Celulose.....	53
5.12 Pré-tratamento de biomassa e sacarificação.....	54
5.13 Fenólicos Totais.....	56
5.14 Ácido Rosmarínico por HPLC.....	57
5.15 Metabolômica.....	58

5.16 Tratamento estatístico.....	60
<b>6. Resultados.....</b>	<b>60</b>
6.1 Similaridade dos genes da via de biossíntese de AR em <i>S. viridis</i> .....	60
6.2 Transformação genética e caracterização molecular.....	61
6.3 Pré-tratamento e Sacarificação de Biomassa.....	65
6.4 Composição da parede celular.....	70
6.5 Composição de lignina por Tioacidólise.....	73
6.6 Fenólicos Totais e Ácido Rosmarínico.....	76
6.7 Biomassa e Inflorescências.....	77
6.8 Metabolômica.....	79
<b>7. Discussão.....</b>	<b>80</b>
7.1 A eficiência de transformação da construção com promotores constitutivos.....	81
7.2 A maquinaria de AR e a presença de <i>TAT</i> em <i>S. viridis</i> .....	83
7.3 Limitações do método de sacarificação.....	84
7.4 Mudanças na composição da lignina e a possível influência do Ácido Rosmarínico na parede celular das plantas transgênicas.....	85
<b>8. Conclusão.....</b>	<b>89</b>
<b>9. Referências.....</b>	<b>91</b>
Anexo I – Representação dos vetores e sequências.....	103
Anexo II – Tabelas.....	109
Anexo III – Cromatogramas.....	117
Anexo IV – Sacarificação de CWR com pré-tratamento de E430.....	120
Anexo V – Expressão gênica de E432 ao longo do crescimento.....	121

## 1. Introdução Geral

O paradoxo envolvendo a crescente demanda energética da sociedade e a dependência de recursos finitos como as fontes fósseis - que atuam como base da sustentação energética atual - segue em pauta nas discussões envolvendo a necessidade de uma transição para novas fontes de energia, que sejam renováveis, e principalmente, menos poluentes. Embora tenha sido parte essencial do desenvolvimento de vários países, principalmente após a primeira e segunda revoluções industriais, a partir de um ponto de vista ambiental, a intensificação do uso de fontes fósseis trouxe consequências graves ao planeta, uma vez que as emissões associadas à queima destes combustíveis acarretam mudanças climáticas significativas, que impactam negativamente diferentes formas de vida no planeta (PENALVA SANTOS *et al.*, 2023).

Em meio à busca por fontes renováveis e não poluentes, surgem oportunidades voltadas à produção de bioenergia fazendo-se uso de substâncias orgânicas, como a biomassa, que é por definição uma matéria orgânica proveniente da fotossíntese de plantas, que pode ser convertida em energia através de processos termoquímicos, bioquímicos, ou ainda, utilizada como matéria-prima para biocombustíveis (Liu *et al.*, 2022). O setor de transportes é responsável por parte significativa das emissões atmosféricas atuais, correspondendo a 25% das emissões totais de CO<sub>2</sub> e 50% do consumo da produção mundial de petróleo (Hortsch; Corvo, 2020). Neste cenário os biocombustíveis desempenham um papel de extrema importância quando se trata de transição energética, pois podem substituir os combustíveis fósseis reduzindo drasticamente as emissões atmosféricas, como por exemplo, o etanol, que tem capacidade de reduzir até 61% das emissões se comparado à gasolina (CHANDEL *et al.*, 2021).

Tratando-se da produção de biocombustíveis, o Brasil ocupa um lugar de destaque, tendo investimentos no setor que remontam à década de 1970, com incentivos a partir do programa Pró-Álcool para a produção de etanol à base de cana-de-açúcar. Tais investimentos aliados à expertise no cultivo de cana-de-açúcar, permitiram que o país ocupasse hoje a segunda colocação em produção de etanol, atrás somente dos Estados Unidos (Robak; Balcerek, 2020). Tanto no Brasil como em outros países, a produção de etanol é baseada na rota industrial que faz uso de um processo microbiológico conhecido como fermentação alcoólica, onde os açúcares sob a forma dos açúcares simples sacarose, glicose e frutose, são convertidos em etanol, energia, CO<sub>2</sub> e outros subprodutos através da ação de leveduras, geralmente *Saccharomyces cerevisiae* (Lopes *et al.*, 2016). Este processo é denominado de primeira geração, e, no Brasil, a sua produção é baseada quase em sua totalidade utilizando-se do caldo

de cana-de-açúcar. Durante o processo de produção do etanol de primeira geração, é gerado um grande volume de bagaço de cana, classificado como biomassa lignocelulósica, material que pode ser utilizado para queima na cogeração de energia, ou ainda na produção do etanol de segunda geração (E2G). Estima-se que no Brasil, até 2030 serão geradas até 26 milhões de toneladas de bagaço de cana (Hortsch; Corvo, 2020). Além do bagaço, outras matérias primas como resíduos agroindustriais e materiais florestais podem ser utilizadas na produção do etanol 2G, todas elas trazendo a vantagem de não competirem com a indústria alimentícia, e ainda estarem disponíveis em grande abundância (AFEDZI; PARAKULSUKSATID, 2023).

Devido às características da biomassa lignocelulósica, a produção do etanol 2G é mais complexa e dispendiosa que o processo de primeira geração, muito por conta da recalcitrância da biomassa, deixando o processo com um baixo custo-benefício. A produção do etanol 2G é baseada em pré-tratamentos, para desestruturar o emaranhado lignocelulósico; hidrólise enzimática, para quebrar os polissacarídeos em açúcares mais simples; e fermentação dos monossacarídeos (Kordala *et al.*, 2024). Diferentes métodos de pré-tratamento podem ser utilizados para quebrar a estrutura lignocelulósica, envolvendo métodos mecânicos, químicos, físico-químicos, e biológicos, sempre com o objetivo principal de remover ou modificar a lignina, reduzir a cristalinidade da celulose, e tornar a hemicelulose mais acessível à hidrólise enzimática (AFEDZI; PARAKULSUKSATID, 2023).

Biomassas de plantas são predominantemente formadas por paredes celulares, um composto formado por várias redes poliméricas, incluindo polissacarídeos e lignina (Loqué; Scheller; Pauly, 2015). A composição de uma biomassa lignocelulósica pode variar de acordo com o tipo de planta, sua idade, condições climáticas e tempo de coleta, porém, em geral, a celulose representa de 40 a 55%, hemicelulose de 24 a 40%, e a lignina de 18 a 25% (Kordala *et al.*, 2024) e ainda há a presença minoritária de pectinas, proteínas, extrativos e cinzas (DALENA *et al.*, 2019).

A lignina é um heteropolímero reticulado amorfo constituído principalmente por três unidades de álcool: coniferílico (guaiacil propanol); cumarílico (*p*-hidroxifenil propanol) e sinapílico (sirigil álcool). Estes monômeros fenólicos são ligados entre si por ligações alquil aril, alquil alquil e aril aril, formando assim uma estrutura tridimensional (Dalena *et al.*, 2019). A lignina está diretamente relacionada à recalcitrância de biomassas lignocelulósicas, uma vez que o polímero de lignina atua como um reforço, impermeabilizando as paredes celulares das plantas e protegendo as microfibrilas de celulose da degradação enzimática (Halpin, 2019). Dessa forma, buscam-se alternativas para aumentar a digestibilidade da biomassa a partir de

estudos em engenharia de lignina. Neste sentido, progressos significativos já foram alcançados na modificação da estrutura, da composição e distribuição da lignina, para reduzir a recalcitrância e aumentar a eficiência de sacarificação sem comprometer o rendimento da biomassa (LOQUÉ; SCHELLER; PAULY, 2015).

A partir da inclusão de monômeros fenólicos compatíveis, é possível modificar características da lignina na parede celular vegetal, tornando-a mais acessível aos pré-tratamentos. Segundo Tobimatsu *et al.*, (2012) perturbações na via dos fenilpropanoides e da via de ácido chiquímico, podem levar a mudanças estruturais massivas no polímero de lignina, a partir de incorporações de intermediários ou outros compostos fenólicos. Neste aspecto, destaca-se o uso do ácido rosmarínico (AR), que tem sua biossíntese justamente ocorrendo através da junção destas duas vias metabólicas, e que já possui enzimas chave bem caracterizadas. Como o AR possui monômeros fenólicos compatíveis com a lignina, além de uma ligação éster facilmente clivável por saponificação, alguns autores sugerem que ele poderia ser incorporado à lignina para que esta facilidade de rompimento possa contribuir no processo de diminuição da recalcitrância da parede celular de biomassas vegetais. Inclusive, na literatura, conforme descrito por Bevilaqua *et al.*, (2019) já há relatos da incorporação de AR na parede celular de plantas vivas resultando em aumento da sacarificação da biomassa pré-tratada com hidróxido de sódio (TOBIMATSU *et al.*, 2012); (VEGA-SÁNCHEZ; RONALD, 2010); (RALPH, 2010); (BEVILAQUA *et al.*, 2019)

Desta forma, este trabalho busca explorar a superexpressão de genes envolvidos na biossíntese de ácido rosmarínico em *Setaria viridis*, uma gramínea de metabolismo C4 utilizada como planta modelo para cana-de-açúcar e outras gramíneas, para verificar se a incorporação deste monômero na lignina pode ter influência positiva na digestibilidade da biomassa. Para a geração das plantas transgênicas foram utilizadas técnicas de biologia molecular e cultura de tecidos, e diversas técnicas de bioquímica também foram aplicadas para verificar a composição da parede celular, mudanças no metabolismo fenólico, bem como a liberação de açúcar ao longo do tempo em comparação com as plantas selvagens.

## 2. Referencial Teórico

### 2.1 Panorama do etanol de segunda geração

O etanol celulósico, classificado como biocombustível de segunda geração tem um grande potencial produtivo, tendo em vista a enorme disponibilidade de matéria-prima para a sua produção. Além da evidente capacidade de substituir o uso do petróleo e dos aspectos econômicos e ambientais, o etanol 2G também traz benefícios do ponto de vista de utilização, permitindo, por exemplo, uma maior potência e uma combustão completa em comparação com a gasolina e o diesel, o que gera como resultado uma menor quantidade de fumaça lançada na atmosfera, contribuindo para minimizar os impactos ambientais. Outros fatores importantes são o alto ponto de fulgor (150°C), que o torna bastante seguro, e as suas boas propriedades de lubrificação, que tem impacto direto na vida útil de motores a combustão (KORDALA *et al.*, 2024).

Por mais que o E2G tenha vários aspectos importantes, a sua produção ainda é extremamente baixa se comparada ao etanol 1G. Em 2023 a produção brasileira de etanol 1G atingiu 32,95 bilhões de litros, enquanto a produção de E2G contribuiu com somente 51 milhões de litros, ou seja, uma participação ínfima. Apesar da produção ainda ser pequena, há grande expectativa de crescimento para o E2G, e isso se reflete no fato de que empresas já investem em novas plantas de produção, como por exemplo a Raízen, uma *joint venture* das empresas Shell e Cosan, com investimentos na ordem de 2,8 bilhões de reais para inauguração de 3 novas plantas de E2G no Brasil (DEGREENIA; WYNNE, 2023).

A valorização do produto também se reflete na ponta final da cadeia, o consumidor. Um estudo realizado por Garcia *et al.*, (2022) evidenciou que consumidores que estão mais familiarizados com o tema de biocombustíveis e energias renováveis têm maior disposição a pagar mais por um combustível composto por gasolina e E2G, valorizando os aspectos ambientais e de sustentabilidade do biocombustível celulósico. Apesar de todos os fatores positivos citados anteriormente, a produção do E2G ainda encontra barreiras a serem superadas para que as projeções se concretizem. Os maiores gargalos estão envolvidos com os custos do processo de produção, principalmente nas duas etapas adicionais do E2G: o pré-tratamento de biomassa e a hidrólise enzimática da celulose e hemicelulose.

O pré-tratamento é responsável por afetar a estrutura integrada da biomassa lignocelulósica, removendo a lignina, para aumentar a acessibilidade e a biodegradabilidade da

celulose e hemicelulose antes das etapas de hidrólise e fermentação. Na literatura, o processo de pré-tratamento vem sendo destacado como uma das etapas mais cruciais do E2G, podendo afetar a viabilidade e a sustentabilidade do processo, devido ao seu grau de complexidade em termos de capital, alto uso de energia e uso de produtos químicos. Alguns autores citam inclusive que o custo somente da etapa de pré-tratamento poderia ter o mesmo valor do custo total do processo de produção, incluindo hidrólise, fermentação e purificação (SOLTANIAN *et al.*, 2020); (DA SILVA; TORRES ORTEGA; RONG, 2016).

A hidrólise é a etapa subsequente do processo após o pré-tratamento, onde a celulose e a hemicelulose são quebradas em açúcares mais simples, grande parte glicose e xilose. A hidrólise geralmente é realizada a partir de coquetéis enzimáticos em um processo ambientalmente amigável que faz o uso de enzimas celulases e hemicelulases que decompõem os polissacarídeos em açúcares fermentescíveis. O sucesso da hidrólise depende muito do quão efetivo foi o pré-tratamento e também da escolha correta do coquetel enzimático de acordo com o tipo de matéria-prima (KORDALA *et al.*, 2024).

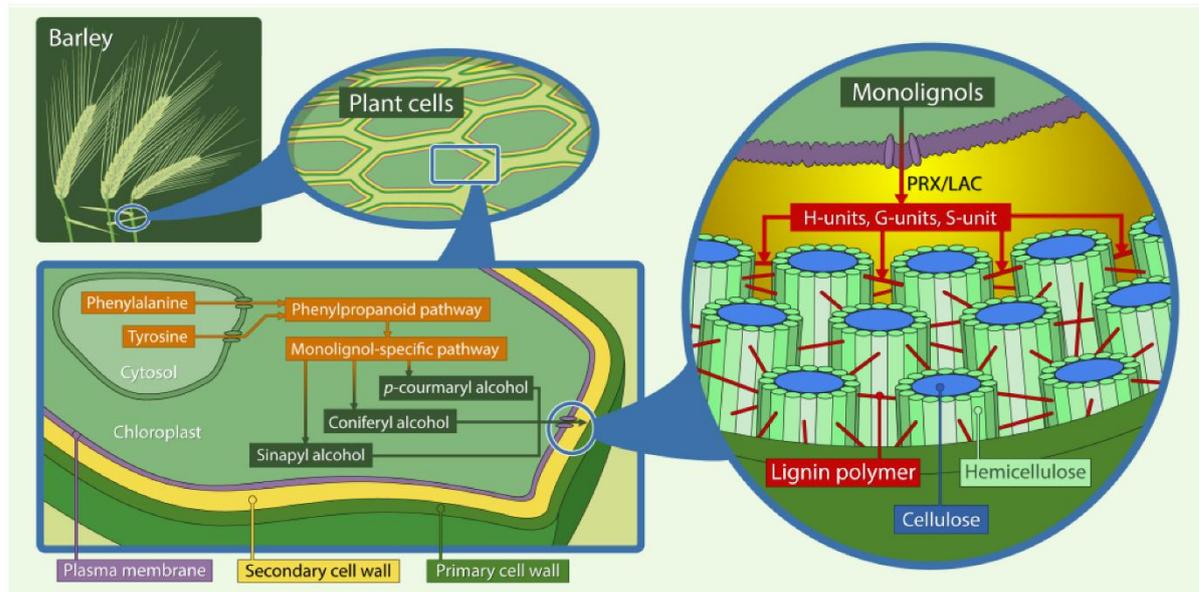
Dessa forma, torna-se importante buscar soluções para aumentar a eficiência da etapa de pré-tratamento, a partir de alternativas que permitam reduzir os custos energéticos e que também permitam que a hidrólise enzimática também ocorra de maneira mais efetiva.

## **2.2 Composição da parede celular de plantas**

Nas plantas, a parede celular é composta principalmente por polissacarídeos, como celulose, hemicelulose, pectinas, e ainda, proteínas. A parede celular é dividida entre a parede celular primária e a parede celular secundária. A parede primária é mais flexível e dinâmica, permitindo a expansão e o crescimento celular, enquanto a parede celular secundária, que deposita-se após o término da expansão da célula em células maduras ou especializadas como o xilema, é mais espessa e rígida, com várias camadas de fibras de celulose, e é importante para a rigidez e suporte mecânico das plantas, principalmente devido a presença da lignina, juntamente com a celulose e a hemicelulose (VOINICIUC; PAULY; USADEL, 2018); (ZHANG *et al.*, 2021).

A composição da parede celular de gramíneas pode ser vista na Figura 1 a seguir.

**Figura 1** – Matriz lignocelulósica da parede celular secundária de gramíneas.



**Nota:** A síntese dos monolignóis ocorre no citosol a partir da fenilalanina e, no caso das gramíneas, também da tirosina, por meio da via dos fenilpropanoides e de rotas específicas para os monolignóis. Posteriormente, essas moléculas são transportadas através da membrana plasmática para a parede celular secundária, onde são oxidadas por peroxidases ligadas à parede celular (PRX) e lacases (LAC), culminando na polimerização e formação da lignina.

**Fonte:** CHRISTENSEN; RASMUSSEN (2019).

Em plantas dicotiledôneas, a parede celular tem uma distribuição mais uniforme dos componentes da parede, o que reflete em uma maior flexibilidade e adaptabilidade, contendo maior quantidade de pectinas, enquanto a parede celular de plantas monocotiledôneas, como *S. viridis*, e gramíneas em geral, tem uma parede celular sob o formato de uma rede mais densa de arabinoxilano e celulose. Em dicotiledôneas, a estrutura mais aberta da parede celular facilita processos de crescimento celular, ao contrário de plantas monocotiledôneas, onde a parede é mais compacta, oferecendo maior resistência a estresses ambientais e também a processos químicos, como a fermentação para produção de etanol (HAAS; PEAUCELLE, 2021).

Em comparação, plantas de metabolismo C3 dicotiledôneas têm parede celular do tipo I, composta principalmente por pectinas abundantes, proteínas estruturais e microfibrilas de celulose que são interligadas por hemiceluloses, como xiloglucanos, formando uma rede. Para gramíneas, como *S. viridis*, *Sorghum bicolor* e *Saccharum spp.* de metabolismo C4, apresentam parede celular do tipo II, com menor quantidade de pectinas, e maior quantidade de

hemiceluloses, como glucuronoarabinosilanas e arabinosilano, além de serem mais rígidas devido a presença da lignina (Zhang; Zhou; Zhang, 2024). A celulose é o principal componente da biomassa lignocelulósica e da parede celular, podendo representar de 40 a 55% da composição. As microfibrilas de celulose são uma estrutura em formato de cabos alongados, composta por 24 a 36 cadeias contendo de 500 a 14.000 moléculas de glicose conectadas por ligações do tipo  $\beta$ -(1,4) presentes tanto na parede celular primária quanto na parede celular secundária. As microfibrilas tem características de serem longas e inelásticas, envolvendo as células em camadas sobrepostas, fornecendo resistência a pressões osmóticas e enrijecendo a parede celular, proporcionando a capacidade de crescimento vertical das plantas (Somerville, 2006); (Kordala *et al.*, 2024); (Meng; Ragauskas, 2014). A hemicelulose representa de 24 a 40% da parede celular de plantas, e em gramíneas é composta primariamente de xilano e  $\beta$ -glucano, sendo o xilano uma cadeia linear de resíduos  $\beta$ -(1,4)-D-xilosil. Além disso, os resíduos de arabinofuranose podem ser ligados à estrutura do xilano formando arabinosilano, que por sua vez pode ser substituído por ácido ferúlico (FA) ou *p*-cumárico (pCA). O FA, produzido na via dos fenilpropanoides, liga-se à lignina de forma covalente, aos polissacarídeos por meio de ligações éster e a proteínas estruturais das paredes celulares de gramíneas através de ligações éter, atuando como um conector universal dos componentes da parede celular (OLIVEIRA *et al.*, 2015); (OLIVEIRA *et al.*, 2020); (KORDALA *et al.*, 2024).

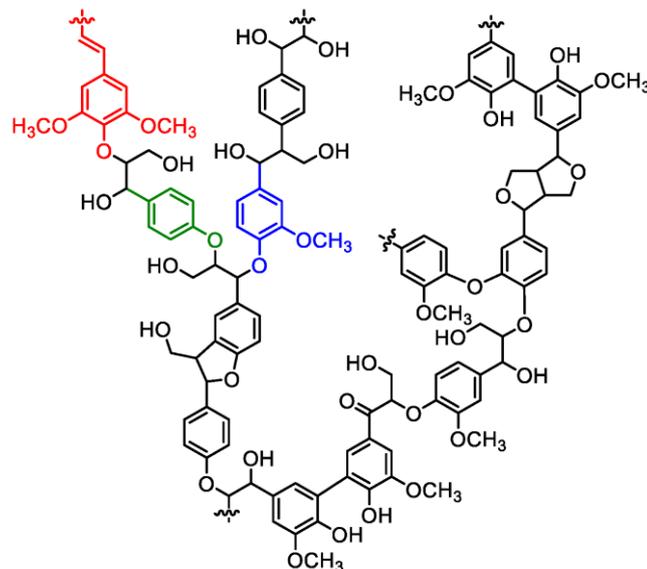
### **2.3 Biossíntese e o papel da lignina na parede celular**

Lignina é um termo que define um grande grupo de polímeros aromáticos resultantes do acoplamento combinatório oxidativo de 4-hidroxifenilpropanoides. Depositados predominantemente nas paredes das células, a lignina tem a capacidade de torná-las rígidas e impermeáveis. A sua biossíntese também pode ser dirigida por fatores de estresse biótico e abiótico, como perturbações na estrutura da parede celular, infecção por patógenos e estresse metabólico (Vanholme *et al.*, 2010). Além do seu papel fundamental no crescimento e desenvolvimento das plantas, protegendo os polissacarídeos de degradação microbiana e conferindo resistência ao apodrecimento, a lignina também desempenha um papel central na recalcitrância da biomassa lignocelulósica, atuando como uma barreira para a sacarificação eficiente da biomassa na produção de biocombustíveis. A remoção da lignina é um processo dispendioso, sendo necessário fazer uso dos pré-tratamentos, que podem remover grande parte da lignina aumentando a eficiência da hidrólise dos polissacarídeos. Também há esforços de

pesquisa para desenvolvimento de plantas com menor grau de deposição de lignina, ou com lignina modificada, que seja mais suscetível à degradação química (SUN *et al.*, 2016); (JÖNSSON; MARTÍN, 2016); (VANHOLME *et al.*, 2010).

A lignina é comumente apontada como sendo composta por três unidades monoméricas principais: *p*-hidroxifenil (H), guaiacil (G) e siringil (S) – Figura 2, que são derivados da polimerização dos álcoois hidroxicinamil, *p*-cumaril, coniferil e álcool sinapil. Contudo, a lignina também é derivada de vários outros monômeros aromáticos. Foram observados vestígios de aldeídos hidroxicinamil, acetatos, *p*-cumaratos, *p*-hidroxibenzoatos e ferulato de tiramina, e também alguns fenólicos agora são reconhecidos como monômeros de lignina. Na literatura, Vanholme *et al.*, (2019) reporta que metabólitos de 11 classes diferentes foram encontrados como sendo de ocorrência natural em monômeros de lignina, incluindo álcoois hidroxicinamil, hidroxiarilpropanóis, ésteres hidroxicinamil, hidroxicinamaldeídos, ácidos hidroxicinâmicos, ésteres de hidroxicinamatos, hidroxicinamidas, hidroxibenzaldeídos, ácidos hidroxibenzóicos, hidroxistilbenos, flavona e tricina. Tais monolignóis podem apresentar abundâncias variadas em diferentes camadas da parede celular, tecidos vegetais e variam com o tipo de planta. A complexidade da macromolécula de lignina tem muito a ver com a variedade de monômeros e a aleatoriedade das ligações que podem ocorrer entre eles (BOERJAN; RALPH; BAUCHER, 2003); (VANHOLME *et al.*, 2019); (HALPIN, 2019).

**Figura 2** – Estrutura da macromolécula de lignina.



**Nota:** Monolignóis majoritários representados em: vermelho, como álcool sinapílico; verde, como álcool *p*-cumarílico; azul, como álcool coniferílico.

**Fonte:** KARUNARATHNA; SMITH (2020).

A rota geral de biossíntese de lignina em plantas monocotiledôneas é evidenciada na Figura 3. Diferentemente de plantas dicotiledôneas, em gramíneas monocotiledôneas existe a via alternativa da deaminação realizada pela enzima TAL – *tirosina amino liase*, onde a tirosina é diretamente convertida em *p*-cumarato. A lignina é sintetizada na parede celular, por meio da polimerização dos monolignóis e todos os blocos de construções conhecidos que fazem parte da biossíntese do polímero são produzidos ou derivados da via geral dos fenilpropanoides. A fenilalanina, derivada da via do shikimato, atua como substrato para a via em todas as plantas, enquanto a tirosina é um substrato adicional em gramíneas. A biossíntese dos monolignóis envolve sucessivas reações de hidroxilação do anel aromático, seguidas por metilações fenólicas e por fim a conversão da cadeia lateral carboxila em um grupamento álcool. (BOERJAN; RALPH; BAUCHER, 2003); (VANHOLME *et al.*, 2013).

Depois de várias etapas envolvendo deaminação, hidroxilação, metilação e redução, os monômeros de lignina são produzidos no citosol e transportados até o apoplasto. Por fim, ocorre a polimerização das três unidades monolignóis principais (G, H e S) por meio da ação de peroxidases (POD) e lacases (LAC) na parede celular secundária. As principais enzimas envolvidas na rota biossintética de lignina são PAL, fenilalanina amônia liase; TAL, tirosina amônia liase; C4H, cinamato 4-hidroxilase; 4CL, 4-cumarato:CoA ligase; CCR, cinamoil-CoA redutase; HCT, hidroxicinamoil-CoA shikimato; C3H, *p*-cumarato 3-hidroxilase; CCoAOMT, cafeoil-CoA O-metiltransferase; F5H, ferulato 5-hidroxilase; CSE, cafeoil shikimato esterase; COMT, ácido cafeico O-metiltransferase; CAD, cinamoil álcool desidrogenase; LAC, lacase; POD, peroxidase (LIU; LUO; ZHENG, 2018); (BOERJAN; RALPH; BAUCHER, 2003).

Para o caso de gramíneas, que são capazes de sintetizar pCA (ácido *p*-cumárico) a partir de fenilalanina ou tirosina, usando fenilalanina/tirosina amônia-liases bifuncionais (PTAL), os metabólitos derivados da tirosina parecem preferencialmente se incorporarem a unidades S e em pCA ligado à parede celular, onde parte deste pCA ligado a parede é ligado a lignina para produzir conjugados de monolignóis da unidade S (Shafiei *et al.*, 2023). Em gramíneas, os monômeros de lignina são sintetizados na via dos fenilpropanoides e de flavona, passam então pela membrana celular e são polimerizados nas paredes celulares secundária. Além disso, a tricina, um flavonoide, é derivada de uma combinação das vias policetílicas de shikimato e acetato-malonato, e foi o primeiro fenólico encontrado em grande parte das gramíneas como sendo um componente fora da via canônica de lignina (Ralph *et al.*, 2019). A via de biossíntese de lignina em gramíneas é demonstrada a seguir.



## 2.4 Engenharia genética da rota de biossíntese de lignina

Modificações na rota de biossíntese de lignina tem sido objeto de estudo de diferentes grupos ao redor do mundo para alterar o conteúdo de lignina de plantas e também a sua composição. Além disso, os estudos permitem entender melhor a função e o papel dos monolignóis, fornecendo informações valiosas sobre aspectos regulatórios da via de biossíntese de lignina, assim como a sua estrutura e papel biológico. Alterações na composição ou estrutura de lignina também podem ser ferramenta importante pensando na diminuição da recalcitrância da biomassa lignocelulósica para a produção de biocombustíveis. Dessa forma, nas últimas duas décadas diversos estudos realizados tiveram como objetivo tentar reduzir o conteúdo de lignina sem afetar o crescimento normal de plantas, e também tornar a lignina mais suscetível aos pré-tratamentos, facilitando o acesso aos açúcares fermentescíveis.

Ainda que plantas com baixas quantidades de lignina possam ser interessantes para a produção de biocombustíveis, mudanças nesta via podem acarretar alterações nas propriedades físico-químicas da lignina, provocando enfraquecimento de vasos, xilema irregular, e isto pode afetar negativamente o crescimento das plantas. Ainda podem ocorrer mudanças mais sutis, como no caso de alteração na relação S/G, que não alteram características visíveis como o crescimento, mas podem alterar o sequestro de moléculas sinalizadoras de defesa da parede celular, e isto faz com que sejam alterados os níveis de transcrição de genes envolvidos com resposta à estresse. Fortes reduções do conteúdo de lignina podem levar a defeitos indesejados graves no crescimento das plantas, além de alterar seus mecanismos de defesa (GALLEGO-GIRALDO *et al.*, 2018); (VANHOLME *et al.*, 2019).

Os genes envolvidos na biossíntese de lignina já foram estudados utilizando mutantes de várias espécies, como *Arabidopsis* - *Arabidopsis thaliana*; milho – *Zea mays*; alfafa – *Medicago sativa*; álamo – *Populus spp.* e tabaco – *Nicotiana tabacum*. Segundo Vanholme *et al.*, (2010), avaliando os estudos realizados, pode-se concluir que a redução da expressão de cada um dos genes da via, tem a capacidade de reduzir os teores de lignina na parede celular em diferentes níveis, a depender do gene, da espécie e do nível de redundância genética.

Um mutante quádruplo de *Arabidopsis thaliana* envolvendo genes *PAL* (*pal1/pal2/pal3/pal4*) teve o seu conteúdo de lignina reduzido se comparado a plantas selvagens, na ordem de 20 a 25%, porém, ao mesmo tempo foram reduzidos os teores de ácido salicílico e a suscetibilidade ao ataque de patógenos aumentou (Huang *et al.*, 2010). A supressão a nível de mRNA do gene de *4CL* em *Pinus radiata* fez com que plantas transgênicas tivessem

o conteúdo de lignina reduzido na ordem de 36 a 50%, porém, observou-se que primariamente devido à depleção da unidade G (guaiacil), a modificação causava um fenótipo de plantas anãs (WAGNER *et al.*, 2009). Os autores Boerjan; Ralph; Baucher (2003) relatam que a regulação negativa de *PAL* e *C4H* em plantas de tabaco reduz grande parte do teor de lignina das plantas. A regulação negativa de *PAL* reduz principalmente unidades G, enquanto a regulação negativa de *C4H* afeta principalmente as unidades S da lignina. Além disso, os autores ainda relatam exemplos onde a regulação negativa de *4CL* reduz o conteúdo de lignina em plantas de tabaco, álamo e *Arabidopsis*, com alterações na unidade S (tabaco), apenas na unidade G (*Arabidopsis*) e sem alterações nos níveis S/G (álamo).

Segundo relata Bewg *et al.*, (2016), utilizando a técnica de RNA de interferência para reduzir a expressão individual de genes envolvidos com a biossíntese de lignina em cana-de-açúcar, foi possível observar que para *F5H*, apesar de não haver redução de lignina, houve aumento de glicose liberada; e para *COMT*, houve uma redução de quantidade de lignina e aumento de liberação de glicose durante hidrólise enzimática. O estudo ainda evidenciou somente pequenas alterações negativas nos parâmetros fenotípicos. Além disso, utilizando a técnica TALEN para induzir mutações na região de *COMT* de cana-de-açúcar, foi possível observar que as mutações neste gene estão diretamente relacionadas à redução do conteúdo de lignina e aumento de sacarificação, sem prejuízo para o desenvolvimento normal da planta (JUNG; ALTPETER, 2016).

A partir destes exemplos podemos concluir que a engenharia da biossíntese da lignina é bastante promissora na manipulação dos teores de lignina em plantas. Isto nos leva à ideia de que as estratégias de engenharia de lignina também podem ser usadas para reduzir a recalcitrância da biomassa lignocelulósica, pensando principalmente na melhoria da eficiência da produção do etanol de segunda geração.

## **2.5 O Ácido Rosmarínico como um monômero de lignina**

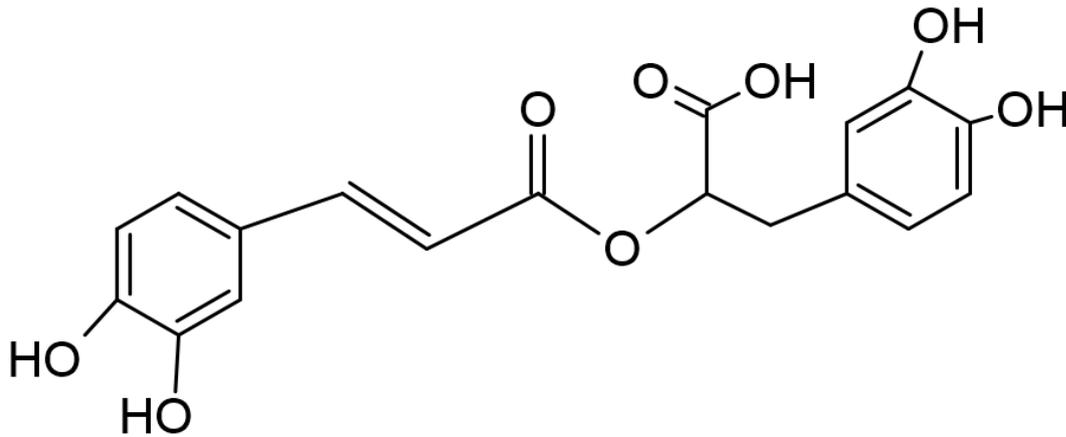
A flexibilidade metabólica e estrutural da biossíntese de lignina, aliado à diversidade de monolignóis que podem ser suas partes componentes, permite que estudos sejam realizados com objetivo de adicionar monômeros e modificar a lignina por rotas alternativas que aumentem a digestibilidade da biomassa lignocelulósica permitindo a remoção da lignina de forma mais fácil (MOTTIAR *et al.*, 2016).

As primeiras evidências da modificação de lignina com este tipo de estratégia foram realizadas através da acilação enzimática de lignina de milho, com unidades *p*-cumarato, e ésteres ferulatos e diferulatos de xilana, também em lignina de milho. Após estudos com modelos miméticos de parede celular e estudos *in planta* provou-se o conceito de que a copolimerização de ésteres ferulados de monolignóis com outros monolignóis poderia facilitar a desestruturação da lignina durante pré-tratamentos químicos. Isto confirmou o valor da engenharia genética de incorporação de monolignóis em lignina como uma das formas de melhorar a deslignificação e a hidrólise enzimática das paredes celulares (MOTTIAR *et al.*, 2016); (GRABBER *et al.*, 2019).

Baseados nestes estudos, Grabber *et al.*, (2019) afirmam que monômeros alternativos que prontamente incorporam à lignina funcionalidades cliváveis de acetil, amida ou éster, podem melhorar a digestibilidade da parede celular após um pré-tratamento levemente ácido ou alcalino. Ainda mais importante, as estratégias de substituição monolignol poderiam evitar fenótipos estruturais ou de desenvolvimento indesejados, associados à regulação negativa de enzimas de biossíntese de lignina, como visto nos exemplos destacados no tópico anterior (VEGA-SÁNCHEZ; RONALD, 2010).

O ácido rosmarínico (AR) é um éster formado pelo ácido caféico e pelo ácido 3,4-dihidroxifenil láctico (Figura 4), e é um constituinte presente em várias famílias de plantas. Recentemente, os genes, enzimas e metabólitos envolvidos na biossíntese de AR foram caracterizados (Weitzel; Petersen, 2011). Segundo Tobimatsu *et al.*, (2012) e Vega-Sánchez; Ronald (2010) o AR é um grande candidato para atuar como uma unidade monolignol de lignina, devido ao fato de possuir uma ligação éster facilmente clivável conectando dois anéis fenólicos para despolimerizar a lignina; e ainda tem a funcionalidade de ácido carboxílico livre, o que o torna capaz de modular a hidrofobicidade da lignina, limitando a ligação com polissacarídeos.

**Figura 4** - Estrutura do Ácido Rosmarínico.



**Fonte:** elaboração própria utilizando ChemSketch (2024).

A rota de biossíntese de AR é complexa, não linear e ocorre em uma rota divergente-convergente. Acontece por meio da via dos fenilpropanoides, e compartilha vários aspectos semelhantes à via de biossíntese de lignina, que foi discutida anteriormente. A enzima PAL, é quem catalisa a conversão da fenilalanina em cinamato. A tirosina é trans aminada pela enzima tirosina aminotransferase (TAT), com 2-oxoglutarato como co-substrato, em 4-hidroxifenilpirúvico (pHPP), que é então reduzido a 4-hidroxifenilático pela enzima hidroxifenil piruvato redutase (HPPR). Tanto NADH como NADPH podem atuar como substratos nesta redução. A C4H catalisa a síntese de *p*-hidroxicinamatos a partir do cinamato formado anteriormente, e a enzima 4CL converte este *p*-cumarato em seu éster de coenzima-A, ativando-o para a reação com malonil-CoA. Na via que leva à formação de AR, a enzima hidroxicinamoil-CoA:hidroxicinamoil transferase (RAS) é quem catalisa a condensação do ácido 4-hidroxifenilático com 4-cumaroil-CoA. Finalmente, estas reações são catalisadas pelos citocromos P450 3-hidroxicinamoil (3-H) e P450 39-hidroxicinamoil (39-H). Posteriormente, o AR pode ser ainda convertido em ácido salvianólico B (SAB) e ácido litospérmico B (LAB) (PETERSEN *et al.*, 2009); (KIM *et al.*, 2013); (TRÓCSÁNYI; GYÖRGY; ZÁMBORINÉ-NÉMETH, 2020).

As enzimas da via geral dos fenilpropanoides PAL, C4H e 4CL são onipresentes nas plantas terrestres, pois são fornecedoras dos precursores para a formação de polímeros de suporte como a lignina e pigmentos UV. A enzima TAT é uma enzima primária que participa também da formação de pHPP, necessário para biossíntese de tocoferóis e plastoquinonas. A via de biossíntese de ácido rosmarínico é evidenciada na Figura 5 (PETERSEN *et al.*, 2009).

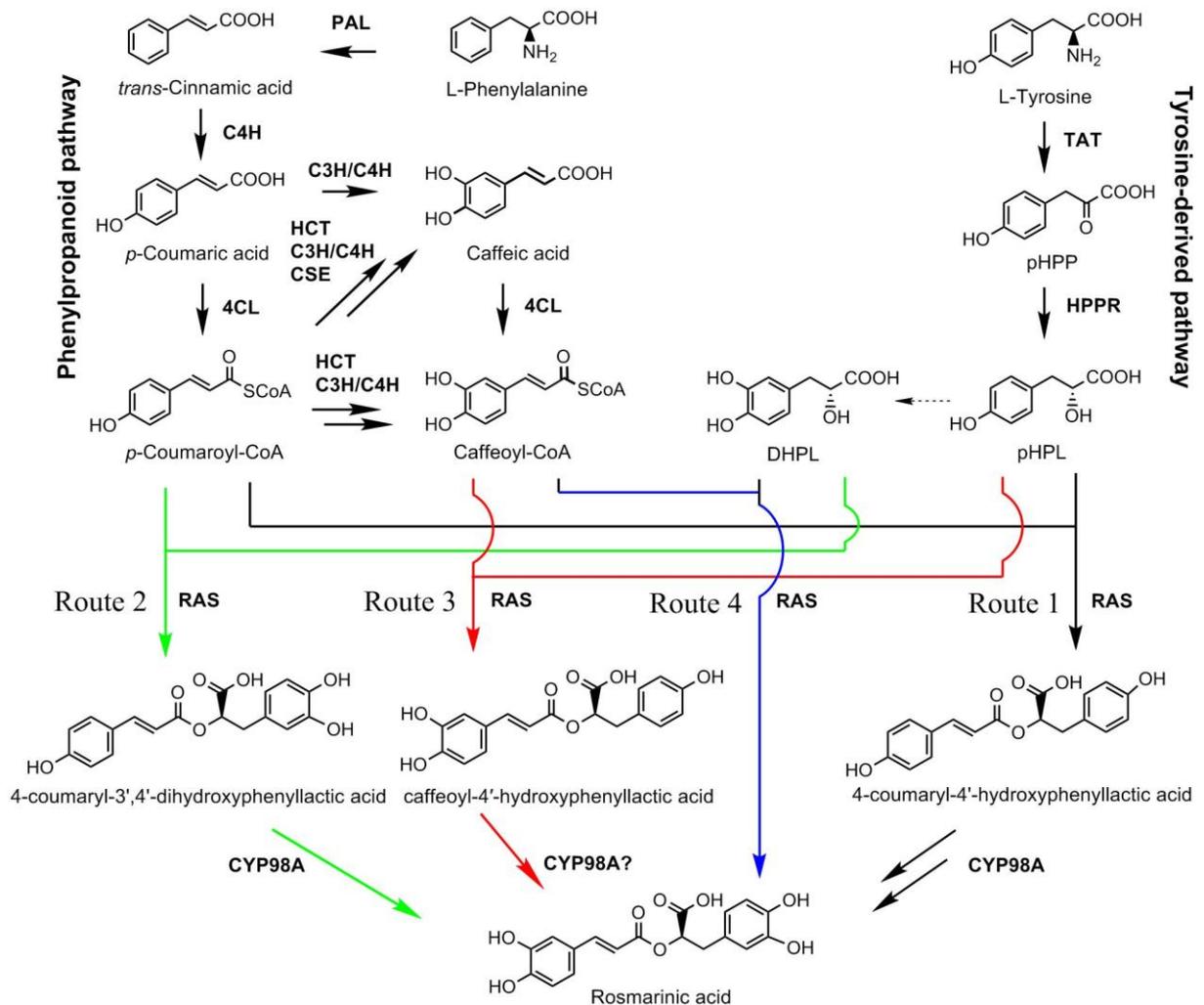
A aplicação exógena de AR foi testada em milho e os resultados sugerem que em uma exposição de curto prazo o AR não altera o conteúdo total de lignina, embora diminua o crescimento das raízes das plantas. Neste caso, foi possível observar alterações na composição monomérica e na atividade das enzimas da rota dos fenilpropanoides, diretamente ligados à biossíntese de lignina. Os autores também relatam que observando a exposição a longo prazo, foi possível verificar que a sacarificação foi melhorada nos colmos, sem nenhum distúrbio na via dos fenilpropanoides (BEVILAQUA *et al.*, 2019).

Um outro estudo relatado na literatura testou a produção de AR em *Salvia miltiorrhiza* através de estratégias de superexpressão de genes envolvidos na biossíntese de AR. Os autores relatam que a superexpressão única dos genes *C4H*, *TAT*, *HPPR*; a superexpressão de *TAT* e *HPPR* em conjunto; e a superexpressão de *HPPD*; aumentaram a produção de AR nas plantas, sendo a maior produção observada naquelas plantas mutantes onde os genes *TAT* e *HPPR* eram superexpressos em conjunto (XIAO *et al.*, 2011).

Além disso, os resultados descritos por Tobimatsu *et al.*, (2012), mostraram que a incorporação de AR à lignina em um modelo mimetizado de parede celular melhora consideravelmente a deslignificação alcalina e a hidrólise subsequente das paredes celulares por preparações de enzimas fúngicas, além de aumentar a degradação da parede celular por enzimas fúngicas e microflora ruminal, mesmo sem pré-tratamentos alcalinos.

Estes resultados indicando o uso de AR como monolignol e a sua incorporação nas plantas a partir da superexpressão de genes forneceram a base teórica para a elaboração da hipótese deste trabalho, onde os genes *TAT* e *HPPR* são superexpressos com a finalidade de modificação da composição e estrutura da lignina, e aumento da digestibilidade da biomassa, utilizando uma planta modelo para a cana-de-açúcar, principal fonte de biomassa para a produção do etanol de segunda geração no Brasil.

**Figura 5 - Rota de biossíntese de Ácido Rosmarínico.**



**Nota:** As setas sólidas representam etapas com um só passo biossintético. Setas duplas representam dois ou mais passos. Setas pontilhadas indicando enzimas cujas reações onde estão envolvidas ainda não foram elucidadas. A rota 1, em preto, foi descrita em *C. blumei* e *P. campanularia*; rota 2 e 3, em verde e vermelho, foram descritas em *S. miltiorrhiza*; e rota 4, foi proposta em *P. vulgaris*.

**Fonte:** Adaptado de Yan *et al.*, (2024).

## 2.6 A presença do Ácido Rosmarínico na família Poaceae e em *Setaria viridis*

Por mais que o AR seja descrito como uma molécula presente em várias famílias de plantas, a sua presença na família Poaceae ainda é muito pouco descrita na literatura. Inicialmente isolado da espécie *Rosmarinus officinalis* (família Lamiaceae), o AR também foi encontrado em outras espécies da família Boraginaceae, sendo apontado como um componente

ativo de várias plantas medicinais, como *Salvia officinalis*, *Mentha x piperita*, *Thymus vulgaris*, *Melissa officinalis*, *Symphytum officinale*; além disso, o AR já foi identificado em três ordens de plantas monocotiledôneas não Poaceae (PETERSEN *et al.*, 2009). Embora seja mais comum das famílias Lamiaceae e Boraginaceae, autores sugerem que AR é um metabólito de distribuição dispersa no reino vegetal, distribuído esporadicamente e com uma complexa evolução da sua rota de biossíntese em plantas terrestres, podendo ter surgido uma vez, e depois se perdido em muitos táxons, ou ainda podendo ter evoluído de forma independente (Levsh *et al.*, 2019). Os autores Guan *et al.*, (2022) relatam a presença de AR em 162 espécies de plantas, em diferentes tecidos vegetais, e destas, 104 são da família Lamiaceae, sendo o gênero *Salvia* o mais comum, com mais de 20 espécies.

Porém, a presença de AR em espécies da família Poaceae ainda é raramente relatada na literatura. O autor Pinto (2015), relatou a identificação de AR a partir de análises de ressonância magnética nuclear de extratos alcoólicos de *Paspalum atratum*, uma gramínea da família Poaceae, mas este foi o único relato encontrado de uma espécie desta família com a presença de AR. Sendo assim, é correto afirmar que a produção de AR em *S. viridis* é incerta, mas não descartada, assim como a presença do seu maquinário de biossíntese.

## **2.7 Utilização de *Setaria viridis* como modelo para cana-de-açúcar**

O uso de plantas modelo para estudos envolvendo transformação genética de culturas importantes já é bem estabelecido na literatura, como por exemplo, estudos em *Arabidopsis thaliana*, um ótimo exemplo de prova de conceito para plantas de metabolismo do tipo C3. Com a crescente busca por alternativas de modificação genética para culturas que são provedoras de biomassa lignocelulósica para a produção de biocombustíveis, surge a necessidade de usar um modelo que traga facilidades do ponto de vista técnico de protocolos de transformação genética e manipulação em pequena escala, uma vez que a cana-de-açúcar tem um genoma extremamente complexo que dificulta o processo de transformação, e aumenta muito o tempo de pesquisa em laboratório.

*Setaria viridis*, da família Poaceae, é uma gramínea de metabolismo do tipo C4, monocotiledônea, pertencente à subfamília Panicoideae, tribo Paniceae, que é irmã da tribo Andropogoneae, um dos grupos de plantas economicamente mais importantes que inclui plantas como cana-de-açúcar, milho e sorgo (Cesarino *et al.*, 2020). Em 2012, o seu genoma foi sequenciado a partir de uma cooperação entre institutos da China e Estados Unidos,

revelando a sua semelhança a nível genômico com outras culturas bioenergéticas, como *Panicum virgatum*. Além disso, *S. viridis* é pertencente à sub-família Panicoidae, juntamente com culturas como sorgo, milho e cana-de-açúcar, todas culturas de metabolismo C4, assim como *S. viridis*. O seu genoma diploide pequeno, a baixa quantidade de DNA repetitivo, e o seu rápido ciclo de vida a tornam um modelo excelente para estudos envolvendo culturas para bioenergia (LATA; GUPTA; PRASAD, 2013).

Pensando justamente neste aspecto onde *S. viridis* pode atuar como planta modelo para estudos envolvendo cana-de-açúcar e outras culturas energéticas, um protocolo de transformação genética baseado em *Agrobacterium tumefaciens* foi desenvolvido por Martins *et al.*, (2015) e estudos envolvendo a transformação bem sucedida de *S. viridis* já foram publicados, como por exemplo De Souza *et al.*, (2018), cujo trabalho evidenciou aumento da digestibilidade da biomassa lignocelulósica a partir da supressão de genes *BAHD* em *S. viridis*. Além disso, é importante destacar que a prova de conceito da utilização de *S. viridis* como planta modelo para cana-de-açúcar já foi relatada na literatura, onde o mesmo grupo de pesquisa citado anteriormente estudou a estratégia utilizando os genes *BAHD* em cana-de-açúcar, obtendo também os mesmos resultados vistos para *S. viridis*, de aumento de digestibilidade da biomassa lignocelulósica (DE SOUZA *et al.*, 2019).

### 3. Justificativa e Hipóteses

A busca por fontes de energia alternativas que possam substituir os combustíveis fósseis é urgente do ponto de vista do setor dos transportes, onde as emissões atmosféricas atingem níveis alarmantes como consequência da queima de gasolina e diesel, principalmente. Fontes alternativas como a biomassa vem ganhando espaço uma vez que a produção de biocombustíveis se mostra capaz de atender as demandas energéticas. Porém, com o aumento crescente da demanda energética, também se buscam alternativas ainda mais sustentáveis para produção de biocombustíveis, surgindo então a possibilidade de aumentar-se a produção do etanol de segunda geração, utilizando biomassa lignocelulósica como matéria-prima.

O Brasil tem a sua produção de etanol baseada principalmente no E1G, e a produção de E2G encontra gargalos técnicos, principalmente voltados para os altos custos das etapas de pré-tratamento e de hidrólise enzimática, muito por conta da presença da lignina no emaranhado lignocelulósico. Alternativas para a modificação da estrutura e da composição da lignina vem ganhando espaço uma vez que este polímero permite substituições em suas unidades

monolignóis, como o ácido rosmarínico, que pode atuar como um monômero alternativo e que que poderia tornar a biomassa lignocelulósica mais suscetível aos pré-tratamentos.

Assim sendo, a principal pergunta deste trabalho é se a superexpressão de genes envolvidos na biossíntese de ácido rosmarínico, através da modificação genética de uma planta modelo para a cana-de-açúcar, denominada *Setaria viridis*, poderia melhorar a digestibilidade da biomassa através do acúmulo de AR na parede celular, sem efeitos deletérios.

A hipótese principal deste trabalho é que a superexpressão dos genes da via de AR poderia facilitar a remoção da lignina por meio de pré-tratamentos, melhorando a digestibilidade da biomassa, aumentando a eficiência de hidrólise enzimática, e consequentemente permitindo reduzir os custos do processo de produção do E2G. Além disso, outras hipóteses foram testadas, para verificar se a expressão aumentada dos genes alteraria a composição da parede celular, principalmente relacionado às quantidades de lignina; verificar se a expressão aumentada dos genes poderia levar a uma maior liberação de açúcar das plantas transgênicas comparadas às plantas selvagens; verificar se a modificação genética poderia causar algum tipo de distúrbio de crescimento, quantidade de biomassa e de sementes nas plantas transgênicas.

#### 4. Objetivos

O objetivo geral deste trabalho é a superexpressão dos genes *SmTAT* e *SvHPPR*, sob o controle de promotores constitutivos e promotores específicos de parede celular secundária em *S. viridis* e a avaliação das características bioquímicas da parede celular por meio de ensaios de composição de parede, hidrólise enzimática e metabolômica.

Colocam-se como objetivos específicos:

- Transformação genética de *S. viridis* utilizando vetores de expressão contendo sequências dos genes e promotores previamente selecionados: *pZmUbi::Smtat::p35Sm::Svhppr* e *pIRX5::Smtat::pIRX1::Svhppr*; para obtenção das linhagens transgênicas;
- Caracterização molecular das plantas geradas a partir de técnicas de biologia molecular, para confirmação de plantas positivas e definição do nível de expressão gênica;

- Ensaios de hidrólise enzimática da biomassa das plantas (sacarificação) para definir as linhagens com maior liberação de açúcar, como *screening* para os demais experimentos;
- Ensaios bioquímicos de caracterização da parede celular: teores de lignina, celulose e hemicelulose; tioacidólise para composição da lignina; purificação de compostos fenólicos para determinação de AR; determinação de fenólicos totais; para verificar se as plantas transgênicas têm composição de parede celular modificada;
- Identificação e quantificação de AR por HPLC e Metabolômica das plantas transgênicas para identificação e quantificação de metabólitos secundários diferenciais envolvidos com o metabolismo de fenólicos, em comparação com plantas selvagens;
- Quantificação de parâmetros de desenvolvimento como biomassa seca total e número de inflorescências das plantas transgênicas.

## 5. Metodologia

### 5.1 Estudo dos genes da maquinaria de AR em *S. viridis* e construção dos vetores de expressão

Para verificar se os genes da rota de biossíntese de AR também estavam presentes em *S. viridis*, procedeu-se para uma análise utilizando a ferramenta BLAST, a partir das sequências gênicas de cada um dos genes da rota: PAL (Phenylalanine ammonia-lyase); C4H (Cinnamate-4-hydroxylase); 4CL (4-Coumarate:CoA ligase); TAT (Tyrosine aminotransferase); HPPD (4-Hydroxyphenylpyruvate dioxygenase); RAS (Rosmarinic acid synthase); C3H (p-Coumaroyl 3'-hydroxylase); HPPR (Hydroxyphenylpyruvate reductase); CYP98A (Cytochrome P450 hydroxylase family). Os identificadores de cada um dos genes foram retirados de referências da literatura. As sequências de cada identificador de cada gene das referências da literatura foram retiradas do GenBank do NCBI para *Salvia miltiorrhiza* e buscados pelo BLAST no Phytozome para *S. viridis* v4.1.

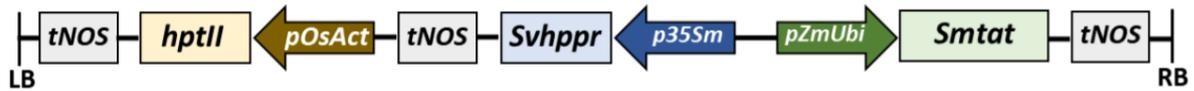
Os vetores de expressão foram construídos a partir de sequências dos genes obtidas no banco de genomas do Phytozome (v.12) disponível em <https://phytozome.jgi.doe.gov/pz/portal.html> e do National Center for Biotechnology Information, NCBI. Duas estratégias diferentes foram escolhidas para a superexpressão dos genes envolvidos na biossíntese de AR. As duas estratégias envolvem a superexpressão de dois

genes, baseados nos resultados do estudo de Xiao *et al.*, (2011), onde a coexpressão destes genes evidenciou maiores quantidades de AR. O primeiro gene *SvHPPR* - *HidroxyPhenylPyruvate Reductase*; identificador Phytozome Sevir.5G009200; e o segundo gene *SmTAT* - *Tyrosine AminoTransferase*; GenBank NCBI: DQ334606.1; cuja sequência foi obtida da anotação de *Salvia mitiorrhiza*. Neste caso, utilizou-se a sequência em *Salvia mitiorrhiza* pois não foram encontrados homólogos para este gene em *S. viridis*, como mostrado ao longo do trabalho. Importante ressaltar que para a construção dos vetores de expressão foram utilizadas as sequências mencionadas, porém, com otimização de códons para monocotiledôneas. Na primeira estratégia, os dois genes foram coexpressos sob o controle de promotores constitutivos *p35Sm* (promotor CaMV 35S modificado para otimização de expressão em monocotiledôneas) e *pZmUbi* (promotor do gene da ubiquitina de milho), para superexpressão de *SvHPPR* e *SmTAT*, respectivamente.

A segunda estratégia de superexpressão dos genes envolveu a expressão espaço-temporal controlada, utilizando-se promotores para expressão na parede celular secundária de *S. viridis*. O gene *SvHPPR* é controlado pelo promotor *pSiIRX1* (1.5 kb à montante do gene *IRX1* de *Setaria italica*; Seita.5G319100) e o gene *SmTAT* atua sob o controle do promotor *AtpIRX5* (1980 pb à montante do start códon do gene *IRX5* de *Arabidopsis thaliana*; At5g44030). Os genes *IRX1* (CesA8) e *IRX5* (CesA4) foram reportados como sendo essenciais para a formação da parede celular secundária em *Arabidopsis thaliana* (HERNÁNDEZ-BLANCO *et al.*, 2007); (GARDINER; TAYLOR; TURNER, 2003). Eles atuam como subunidades da celulose sintase (CesA) e, junto a um terceiro gene, são responsáveis pela síntese de celulose nos vasos do xilema após a expansão celular. Os vetores de expressão são compostos pelo marcador de seleção *hptII* (hygromycin phosphotransferase II), um gene que confere resistência ao antibiótico Higromicina, sob o controle do promotor da actina de arroz (*pOsAct*). Os vetores foram sintetizados pela empresa DNA Cloning Service, sediada na cidade de Hamburgo, Alemanha. Um esquema dos vetores de expressão está apresentado na Figura 6. As sequências dos genes e promotores, além da ilustração completa dos vetores podem ser encontradas no Anexo II – vetores e sequências genéticas. Com a finalidade de facilitar a identificação, os vetores foram identificados como E430, para as o vetor utilizado promotores constitutivos; e E432 para o vetor utilizando promotores de parede celular secundária, assim como as linhagens obtidas para cada um dos vetores utilizados.

**Figura 6** - Esquema de construção dos vetores de expressão.

(a) E430: vetor utilizando promotores constitutivos.



(b) E432: vetor utilizando promotores para expressão espaço temporal controlada na parede celular secundária.



**Fonte:** do autor (2024).

Os vetores foram eluidos a partir da indicação do fornecedor, adicionando-se 50  $\mu\text{L}$  solução TE na membrana do tubo contendo os vetores. Aguardou-se 2 minutos para a absorção completa e centrifugou-se a 14.000  $g$  por 2 minutos a temperatura ambiente. Procedeu-se com a quantificação em espectrofotômetro e eletroforese em gel de agarose, corado com brometo de etídio para verificar a integridade das amostras, em triplicata.

Os vetores de expressão foram inseridos em *E. coli* e em cepas EHA105 de *Agrobacterium tumefaciens* ambos através do método de transformação de choque térmico, descrito por Brasileiro; Carneiro (2015). Foram utilizadas células competentes de *E. coli* previamente preparadas que já estavam disponíveis no laboratório. A transformação por choque térmico ocorreu iniciando-se com a preparação da suspensão bacteriana que estava estocada a  $-80^{\circ}\text{C}$ . O tubo contendo as células competentes foi deixado para descongelamento no gelo, de onde retirou-se uma alíquota de 100  $\mu\text{L}$  para um micro tubo previamente resfriado. Ao DNA plasmidial, que foi previamente ressuspendido, adicionou-se cerca de 20 nanogramas às células competentes, misturou-se cuidadosamente e os microtubos foram deixados no gelo por 30 minutos.

Os tubos então foram transferidos para um banho a  $42^{\circ}\text{C}$ , durante 90 segundos, sem agitação. Após, transferiu-se os tubos rapidamente para o gelo durante 2 minutos. Adicionou-se então 800  $\mu\text{L}$  de meio Luria Bertani (LB). Misturou-se suavemente e os tubos foram transferidos novamente para o banho a  $37^{\circ}\text{C}$  por 45 minutos. Após, utilizou-se de centrifugação em velocidade máxima por 30 segundos para concentrar as células e descartou-se 800  $\mu\text{L}$  de

sobrenadante. Os 100 µL restantes foram então plaqueados em meio LB contendo o antibiótico Espectinomicina 100 mg/L para seleção das bactérias. As placas foram colocadas para crescimento em incubadora a 37°C por 16 horas. Após a transformação, foi realizado um estoque de glicerol de *A. tumefaciens* e também de *E. coli*, com a finalidade de manter um estoque das linhagens trabalhadas no laboratório.

Para a incorporação dos vetores em *A. tumefaciens*, também foi utilizado o protocolo por choque térmico descrito por Brasileiro; Carneiro (2015). Do mesmo modo, utilizou-se de células competentes previamente preparadas disponíveis no laboratório. Preparou-se inicialmente um pré-inóculo, contendo uma colônia isolada de *A. tumefaciens* coletada com uma alça de platina, em um tubo de ensaio, com 3 mL de meio LB e os antibióticos Espectinomicina 100 mg/L e Rifampicina 50 mg/mL. Incubou-se a 28°C sob agitação overnight. Incubou-se a cultura no gelo por 15 minutos. Centrifugou-se as células a 1968 g por 15 minutos a 4°C. Descartou-se o sobrenadante e as células foram ressuspendidas em 1 mL de CaCl<sub>2</sub> 20 mM em glicerol 10%. Retirou-se uma alíquota das células competentes e manteve-se no gelo até descongelar. Adicionou-se 1 µg de DNA plasmidial e misturou-se delicadamente. Incubou-se no gelo por 30 minutos e após, em nitrogênio líquido por 2 minutos. Depois desse período, incubou-se as células a 37°C por 5 minutos. Adicionou-se 1 mL de meio LB, e incubou-se por 2 horas a 28°C. Após isso, espalhou-se com auxílio de alça de Drigalsky 100 µL da suspensão de células em placas contendo meio LB/ágar e os antibióticos Espectinomicina 100 mg/L e Rifampicina 50 mg/L. Incubou-se as placas a 28°C e aguardou-se por 3 dias o crescimento das colônias. Para a confirmação da presença do vetor transferido, realizou-se PCR utilizando as colônias formadas e primers específicos de cada vetor.

Realizou-se ainda um protocolo de miniprep para cada um dos vetores transformados, a partir das instruções do kit GeneJET Plasmid Miniprep Kit, Thermo Scientific. Foi realizado em triplicata. Adicionou-se 250 µL da solução de ressuspensão ao pellet obtido a partir de 1,5 mL de cultura de *Agrobacterium* transformada, em meio LB contendo os antibióticos de seleção. Após, adicionou-se 250 µL de solução de lise, homogeneizou-se de forma vigorosa até que a solução se tornasse viscosa. Adicionou-se após 350 µL de solução de neutralização e imediatamente homogeneizou-se invertendo-se os tubos. Centrifugou-se por 5 minutos a 12.000 g para concentrar o debris celular e DNA cromossomal. Transferiu-se o sobrenadante para a coluna do kit e centrifugou-se a 12.000 g por 1 minuto. Adicionou-se 500 µL da solução de lavagem na coluna e centrifugou-se. Repetiu-se o processo de lavagem outra vez. Transferiu-se a coluna para um novo tubo e adicionou-se 50 µL da solução de eluição no centro da coluna.

Incubou-se a temperatura ambiente por 2 minutos e centrifugou-se a 12.000 g por 2 minutos. Após o processo, as amostras de miniprep foram quantificadas utilizando espectrofotômetro e a correta configuração do equipamento.

## 5.2 Transformação genética de *S. viridis*

A transformação genética de *S. viridis* foi baseada no método apresentado por Martins *et al.*, (2015), com adaptações. Os experimentos foram realizados no laboratório de Evolução e Diversidade, campus São Bernardo do Campo da UFABC. O primeiro passo foi a seleção de sementes maduras de *S. viridis* selvagens, genótipo A10.1, para passarem pelo processo de escarificação, para remoção da casca, sem danificar o embrião, com auxílio de lixa e pinça. A desinfestação das sementes ocorreu em câmara de fluxo laminar, fazendo-se uso de 20 µL de Tween 20 e 1 mL hipoclorito de sódio 10-12%, em um tubo eppendorf de 2 mL, contendo de 500 a 1000 sementes, por um período de 5 minutos com agitações manuais. Após o tempo, as sementes foram lavadas com água destilada autoclavada 5 vezes, e deixadas para secar em papel filtro autoclavado.

Cerca de 20 a 25 sementes foram então transferidas para placas Petri de plástico, contendo meio de cultura CIM (descrição detalhada dos componentes do meio de cultura CIM na Tabela 1, Anexo I), e colocadas para incubação a  $25 \pm 2^\circ\text{C}$ , por 4 semanas sem luz, para que houvesse a indução dos calos embriogênicos. Após a formação dos calos no escuro, a próxima etapa realizada foi o isolamento dos calos embriogênicos, selecionados com ajuda de pinça, bisturi e lupa modelo Leica EZ4, em fluxo laminar. Os calos selecionados foram transferidos novo meio CIM, incubados a  $25 \pm 2^\circ\text{C}$  sem luz por uma semana.

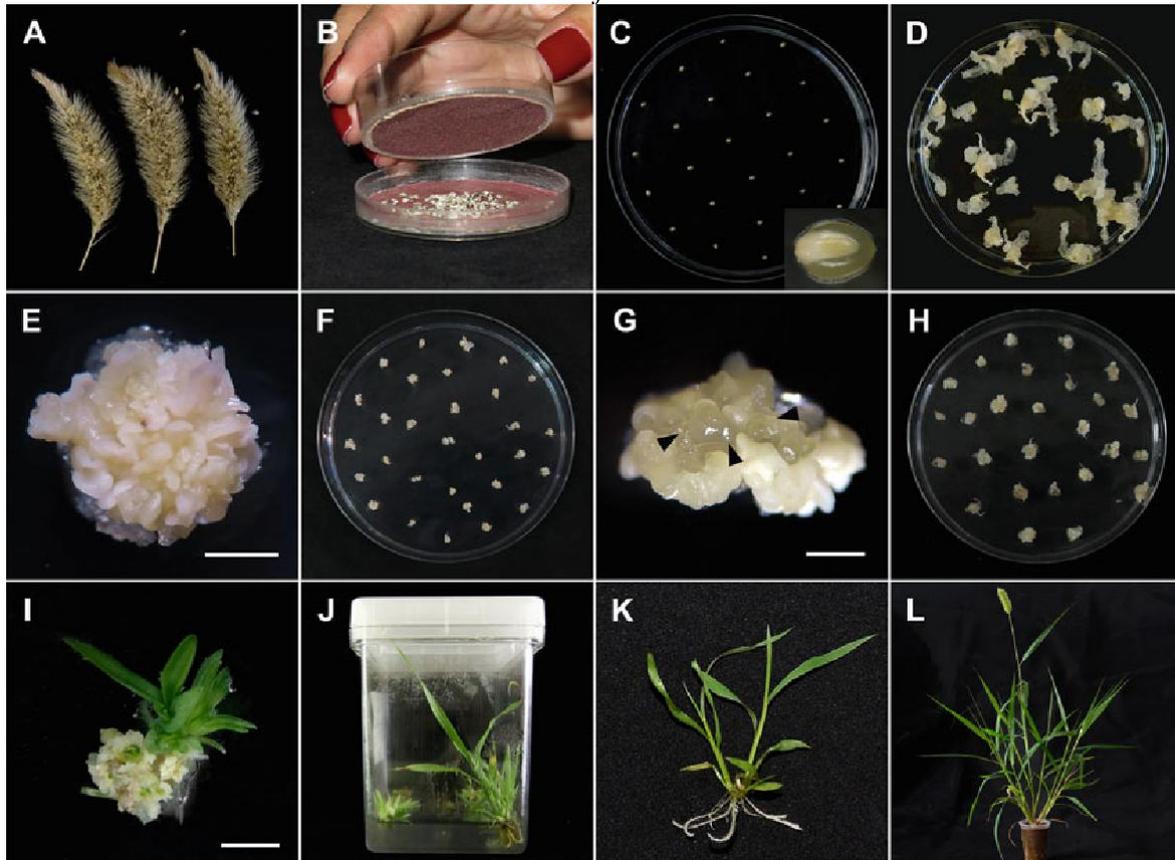
O pré-inóculo da cultura de *A. tumefaciens* foi preparado a partir de colônias da cepa EH105 previamente transformadas e confirmadas, crescidas em meio LB sólido, mantidas em geladeira. Uma colônia foi repicada para um Falcon de 15 mL contendo 3 mL de meio LB e os antibióticos de seleção Espectinomicina 100 mg/L e Rifampicina 50 mg/L, e incubada overnight, com tubo inclinado em incubadora a  $28^\circ\text{C}$  e 180 rpm. No dia seguinte, 2 mL do pré-inóculo foram transferidos para um Erlenmeyer previamente autoclavado de 150 mL. Adicionou-se também outros 18 mL de meio LB contendo Espectinomicina e Rifampicina nas mesmas concentrações descritas anteriormente, além de Acetoseríngona 200 µM. Incubou-se sob agitação a  $28^\circ\text{C}$  até atingir a OD<sub>600</sub> ótima para transformação de 0.6. A OD foi medida utilizando-se cubeta de plástico e 1 mL de suspensão em espectrofotômetro modelo Eppendorf

BioPhotometer Plus. Após, centrifugou-se a 1507 *g* por 15 minutos, descartou-se o sobrenadante e o pellet foi ressuspensionado em 3 mL de meio CIM líquido (Tabela 2 – Anexo I).

Transferiu-se então os calos para um tubo Falcon de 50 mL, e em seguida 10 mL da suspensão bacteriana, juntamente com 100  $\mu$ L de Synperonic (Poloxamer 188 solution - P5556 Sigma). Após 5 minutos de incubação e homogeneização suave, transferiu-se os calos para papel filtro para secagem do excesso de cultura bacteriana, e transferiu-se os calos para placas petri contendo meio de cultura CIM acrescido de Acetoseríngona 200  $\mu$ M. Incubou-se a 22°C sem luz por 3 dias. Após os 3 dias de cocultivo, transferiu-se os calos para meio CIM acrescido de Timentin 150 mg/L; após sete dias, transferiu-se para meio CIM acrescido de Timentin e Higromicina 30 mg/L e incubou-se por mais 7 dias sem luz a 25°C. Após esta etapa, os calos foram transferidos para meio de regeneração de raízes MRS (Tabela 3 – Anexo I) acrescido de Timentin 150 mg/L e Higromicina 30 mg/L. Incubou-se as placas a  $26 \pm 2^\circ\text{C}$  em luz  $150 \mu\text{mol m}^{-2}\text{s}^{-1}$  com fotoperíodo de 16 horas luz/8 horas escuro. Os explantes que demonstraram crescimento de raízes após 20 dias, foram transferidos para um novo meio MRS. Após confirmação de que o explante é um bom candidato, realizou-se a transferência para o meio de desenvolvimento, em frascos de vidro (meio MD – Tabela 4 – Anexo I), e mantinham-se a  $26 \pm 2^\circ\text{C}$  em luz  $150 \mu\text{mol m}^{-2}\text{s}^{-1}$  com fotoperíodo de 16 horas luz/8 horas escuro.

Após crescimento da parte aérea, transferiu-se o explante candidato para copos plásticos contendo substrato:vermiculita em razão 3:1, acrescido de NPK. Os candidatos foram mantidos em casa de vegetação, com controle de temperatura a  $26 \pm 2^\circ\text{C}$ , fotoperíodo 16/8 horas (luz/escuro) com intensidade de luz de  $500 \mu\text{mol m}^{-2}\text{s}^{-1}$ , e eram regados diariamente. A Figura 7 evidencia todas as etapas desde a escarificação das sementes até a aclimação do explante candidato.

**Figura 7** - Etapas do processo de transformação genética de *S. viridis* por *Agrobacterium tumefaciens*.



**Nota:** (A) Separação das sementes; (B) escarificação; (C) distribuição em meio CIM; (D) indução de calos; (E) calo embriogênico; (F) calos embriogênicos isolados; (G) calo embriogênico translúcido; (H) calos embriogênicos após transformação; (I) formação de parte aérea; (J) formação de parte aérea e raízes em meio de desenvolvimento; (K) explante candidato a aclimação; (L) evento aclimatado.

**Fonte:** MARTINS *et al.*, (2018).

### 5.3 Extração de DNA e confirmação via PCR

Após um período de 2 semanas de crescimento das plantas T0 aclimatadas, procedeu-se para a confirmação da integração do vetor de expressão no genoma da planta, a partir de extração de DNA, seguido de PCR (*Polymerase Chain Reaction* – Reação em Cadeia da Polimerase) utilizando primers específicos e eletroforese em gel de agarose. Para a extração do DNA das plantas, utilizou-se um protocolo baseado no método CTAB (Brometo de Cetilmetilamônio). As etapas do processo de extração de DNA são descritas a seguir. Iniciou-se com a coleta dos tecidos vegetais, a partir do uso de tubos eppendorf de 2 mL, tesoura esterilizada, e nitrogênio líquido em isopor. Após higienização da tesoura e das mãos protegidas por luvas com álcool 70%-100%, fez-se a coleta de cerca de 7 a 10 cm de folhas maduras,

verdes, imediatamente inseridas nos tubos eppendorf e no nitrogênio líquido. A etapa de maceração do material foi realizada com auxílio de duas *beads* magnéticas por tubo eppendorf, e equipamento disruptor de células Loccus L-Beader 6. Acrescentou-se ao tubo eppendorf contendo o tecido vegetal macerado 700  $\mu\text{L}$  de CTAB comercial, e 1,4  $\mu\text{L}$  de  $\beta$ -mercaptoetanol, homogeneizou-se e colocou-se em termobloco modelo Ciencor Agimax a 65°C por 30 minutos e homogeneizações manuais a cada 10 minutos. Após, acrescentou-se 700  $\mu\text{L}$  de solução clorofórmio:álcool isoamílico 24:1, e misturou-se manualmente por 10 minutos. Procedeu-se para a centrifugação a 3075  $g$  por 10 minutos, remoção do sobrenadante para um novo tubo, e repetiu-se o processo até nova centrifugação. A partir do novo sobrenadante, acrescentou-se 1  $\mu\text{L}$  de RNase A (100 mg/L) para cada 100  $\mu\text{L}$  de volume resgatado. Incubou-se no termobloco a 37°C por 30 minutos. Adicionou-se 0,6 volume de isopropanol gelado, e centrifugou-se a 12298  $g$  por 20 minutos. O sobrenadante foi descartado e o pellet lavado com etanol 70%, seguido de centrifugação a 12298  $g$  por 3 minutos, por duas vezes. O pellet foi ressuspenso com água ultrapura previamente autoclavada. A concentração das amostras foi determinada a partir de espectrofotômetro do tipo BioDrop uLite+, utilizando-se 1  $\mu\text{L}$  de amostra e a água ultrapura como branco. Ainda se realizou uma corrida de eletroforese em gel de agarose somente para verificar a integridade das amostras de DNA, utilizando-se de 2  $\mu\text{L}$  de amostra, 3  $\mu\text{L}$  de água ultrapura autoclavada e 2  $\mu\text{L}$  de *loading buffer*.

Para a confirmação das plantas utilizou-se o DNA extraído e primers específicos em reações de PCR. O kit de PCR da Sinapse Biotecnologia foi utilizado. Otimizou-se as concentrações sugeridas pelo fabricante para reações individuais de 25  $\mu\text{L}$ , como segue:

- Primers Forward 10 mM: 1  $\mu\text{L}$ ;
- Primer Reverse 10 mM: 1  $\mu\text{L}$ ;
- PCR Buffer/ Tampão Taq: 2,5  $\mu\text{L}$ ;
- dNTP's: 0,5  $\mu\text{L}$ ;
- $\text{MgCl}_2$ : 2  $\mu\text{L}$ ;
- Taq Polimerase: 0,25  $\mu\text{L}$ ;
- Template de DNA: 1-10  $\mu\text{L}$ ;
- Água ultrapura autoclavada: Completar para 25  $\mu\text{L}$ .

O termociclador modelo ProFlex da marca Applied Biosystems foi utilizado em reações de 30 a 35 ciclos:

- Desnaturação (1x): 95°C por 5 minutos;
- Anelamento (30 a 35x): 94°C por 30 segundos, 55-68°C por 30 segundos e 72°C por 1 minuto;
- Extensão (1x): 72°C por 10 minutos.

As sequências dos primers utilizados nas reações de PCR para confirmação de plantas positivas estão descritos na Tabela 5 do Anexo I. Utilizou-se primers para identificação da sequência de resistência ao antibiótico Higromicina, denominado *hptII*; e para identificação da sequência do promotor de Actina do gene de resistência a Higromicina, denominado *pAct*. Também foi utilizado o par de primers denominado *CAC* para identificação de uma sequência comum em *S. viridis*, para evidenciar integridade de amostras de DNA. A escolha do par de primers a ser utilizado baseou-se na eficiência das reações de PCR. Para verificar os resultados dos experimentos de PCR, utilizou-se eletroforese em gel de agarose, na concentração de 1%, e também marcador de peso molecular e *loading buffer* da marca Sinapse Biotecnologia, em corridas que variavam entre 30 minutos e 2 horas, a 120 volts usando uma cuba do tipo Mupid-EXu da marca Advance. Os experimentos foram realizados no laboratório de Evolução e Diversidade, campus São Bernardo do Campo da UFABC.

#### **5.4 Avanço de geração para obtenção de plantas homozigotas**

Para obtenção de plantas homozigotas, ou seja, aquelas que apresentam a presença do transgene inserido em ambos os alelos, utilizou-se da lei de Mendel para obtenção de plantas em segregação 3:1, em um primeiro momento. Os experimentos foram realizados no laboratório de Evolução e Diversidade, no campus São Bernardo do Campo da UFABC. Estas plantas eram denominadas T1 – segunda geração. As sementes destas plantas, eram então denominadas de T2 – terceira geração, e eram submetidas ao teste de homozigose. Para o avanço da geração das plantas, utilizou-se as sementes coletadas após a senescência das plantas T0 – primeira geração, confirmadas por PCR. Selecionou-se 100 sementes maduras, que foram submetidas à quebra de dormência utilizando 1 mL de ácido sulfúrico 100% por 15 minutos, e posterior lavagem de 5 a 7 vezes com água destilada.

Após, em câmara de fluxo laminar, submeteu-se as sementes à desinfestação utilizando-se 1 mL de Hipoclorito 10-12% e 20 µL de Tween 20 por 5 minutos com agitações manuais e posterior lavagem com água destilada autoclavada. Após secagem em papel filtro, as transferiu-se as sementes para placa de petri de vidro meia borda alta, contendo meio de cultura ½ MS (Tabela 6 – Anexo I), suplementado com o antibiótico de seleção de plantas Higromicina 60 mg/L. Também se acrescentou de 10 a 20 sementes WT para controle. As placas foram incubadas em sala de cultura de tecidos à temperatura de  $26 \pm 2^\circ\text{C}$ , intensidade luminosa de  $150 \mu\text{mol m}^{-2}\text{s}^{-1}$  e fotoperíodo de 16 horas luz/8 horas escuro.

Após 10 a 15 dias de crescimento, avaliou-se o número de plantas com raízes e parte aérea, observando o teste estatístico do qui-quadrado de Pearson para comparar proporções, ou seja, possíveis divergências entre as frequências observadas e esperadas de um determinado evento, e verificar se as plantas cumpriam a segregação do tipo 3:1. Para *S. viridis* espera-se uma proporção de 3 plantas tolerantes para 1 planta suscetível, seguindo as regras segregação mendeliana. Para o cálculo observou-se a seguinte fórmula de acordo com a Equação 1:

#### **Equação 1 – Cálculo do qui-quadrado**

$$x = \frac{(\text{Plantas observadas tolerantes} - \text{Plantas esperadas tolerantes})^2}{\text{Plantas esperadas tolerantes}} + \frac{(\text{Plantas observadas suscetíveis} - \text{Plantas esperadas suscetíveis})^2}{\text{Plantas esperadas suscetíveis}}$$

A partir do número total de sementes definiu-se os valores de plantas esperadas tolerantes e suscetíveis; e a partir do número total de sementes que germinaram, sendo tolerantes ou suscetíveis, definiu-se os valores correspondentes. Para avaliar o resultado do teste, utilizou-se um teste de hipótese, utilizando-se um valor de  $x^2$  tabelado em uma tabela de distribuição *t* de Student, fixando-se o nível de significância em 5% ou  $p = 0,05$  e um grau de liberdade de 1.

Plantas que apresentaram segregação 3:1 foram transferidas para substrato:vermiculita (3:1) contendo NPK, e colocadas em casa de vegetação para crescimento. Tais plantas foram denominadas T1 e também passaram pelo processo de confirmação por PCR utilizando primers específicos. Após a confirmação, aguardou-se a senescência para coleta das sementes e avanço para a geração T2, e ainda para verificar qual ou quais plantas eram homozigotas.

Para verificar a homozigose das plantas, procedeu-se para quebra de dormência, desinfestação e crescimento em placa ½ MS contendo Higromicina 60 mg/L, da mesma forma descrita anteriormente. Neste caso, após 10 a 15 dias de crescimento, observou-se placas onde o crescimento das plantas atingisse valores próximos a 100% para ser considerada uma planta homozigota. Assim, as plantas também foram transferidas para copos plásticos contendo substrato:vermiculita (3:1) e NPK, e colocadas para crescimento em casa de vegetação.

### **5.5 Extração e tratamento de RNA, síntese de cDNA e PCR quantitativa em tempo real**

Uma vez que as plantas homozigotas foram confirmadas, procedeu-se para a avaliação do nível de expressão dos genes presentes no vetor de expressão, para verificar se realmente havia superexpressão de *SmTAT* e *SvHPPR* nas plantas transgênicas em comparação com as plantas selvagens. Para isto, primeiramente realizou-se a extração de RNA das plantas, seguindo protocolo do kit RNazol<sup>®</sup> RT da Sigma. A coleta do tecido vegetal e preparação do tecido foram realizadas da mesma forma que descrito anteriormente para DNA. Uma vez com o tecido macerado, adicionou-se 1 mL de RNazol e homogeneizou-se em vórtex. Centrifugou-se a 12.000 g por 5 minutos. Transferiu-se o sobrenadante para um novo tubo e adicionou-se 400 µL de água ultrapura autoclavada, homogeneizou-se por 15 segundos e manteve-se em temperatura ambiente por 15 minutos. Centrifugou-se a mistura a 12.000 g por 15 minutos e transferiu-se o sobrenadante para um novo tubo.

Adicionou-se o mesmo volume de isopropanol 100% e manteve-se na bancada por 10 minutos. Centrifugou-se a 12.000 g por 10 minutos, descartou-se o sobrenadante, e procedeu-se para a lavagem do pellet com 500 µL de etanol 75% por duas vezes. Centrifugou-se novamente, a 6.000 g por 3 minutos. Ressuspendeu-se o pellet em água ultrapura autoclavada e fez-se uso do vórtex para homogeneização. Para quantificação das amostras utilizou-se do mesmo equipamento descrito anteriormente para DNA, porém com as especificações corretas para amostras de RNA. Também se realizou uma corrida de eletroforese em gel de agarose para verificar a integridade das amostras, da mesma forma que para DNA. Procedeu-se então para o tratamento das amostras de RNA utilizando-se do kit RQ1 RNase-Free DNase, Promega. Segundo as instruções do fabricante, preparou-se um mix de DNase e DNase Buffer 1:1 e adicionou-se 2 µL em cada tubo. A quantidade amostra a ser adicionada foi calculada a partir das quantificações das amostras de RNA. E adicionou-se ainda água ultrapura autoclavada para

completar a reação de 8  $\mu\text{L}$ . No termociclador, foi programado um ciclo de ciclo de 30 minutos a 37°C, seguido de 5 minutos a 4°C e após, 10 minutos a 65°C. Durante o ciclo de 5 minutos, abriu-se o termociclador e acrescentou-se 1  $\mu\text{L}$  de *stop solution* em cada um dos tubos. Para a síntese de cDNA utilizou-se o kit Revert Aid First Strand cDNA Synthesis da ThermoScientific. De acordo com o protocolo, foram realizados os cálculos para que a reação ocorresse com 1 ng de amostra de RNA tratado. Adicionou-se a quantidade necessária de amostra, e completou-se com água ultrapura autoclavada para uma reação final de 12  $\mu\text{L}$ . Preparou-se um mix contendo os seguintes componentes:

- 5x Reaction Buffer: 4  $\mu\text{L}$ ;
- RiboLock RNase Inhibitor: 1  $\mu\text{L}$ ;
- 10 mM dNTP Mix: 2  $\mu\text{L}$ ;
- RevertAid M-MuLV RT: 1  $\mu\text{L}$ .

Totalizando 20  $\mu\text{L}$ , os tubos foram colocados em termociclador a 42°C por 60 minutos, e após a 70°C por 5 minutos. O cDNA sintetizado foi armazenado em freezer -80°C.

Para a reação de PCR quantitativa em tempo real, utilizou-se o kit Platinum<sup>®</sup> SYBR<sup>®</sup> Green qPCR SuperMix-UDG with ROX (5  $\mu\text{L}$  de SYBR; 0,2  $\mu\text{L}$  para cada primer em concentração 10 mM; 3,6  $\mu\text{L}$  de água ultrapura autoclavada; e 1  $\mu\text{L}$  de *template* em uma diluição 1:10). Utilizou-se o equipamento da Applied Biosystems, que faz uso do software StepOne 2.3 para geração das curvas e dados de ct. As sequências dos primers utilizados nos experimentos de qPCR estão descritos na Tabela 7 do Anexo I. O ciclo de reação foi programado para 50°C por 2 minutos; 95°C por 2 minutos; e 40 ciclos de 95°C por 15 segundos, seguidos de 60°C por 30 segundos. Os cálculos para determinação do  $\Delta\Delta\text{CT}$  foram realizados de acordo com a fórmula proposta por Schmittgen; Livak (2008), mostrada abaixo como Equação 2:

### **Equação 2 – Cálculo da expressão gênica pelo $\Delta\Delta\text{CT}$**

$$2^{-\Delta\Delta\text{Ct}} = \frac{[(\text{Ct do gene de interesse} - \text{Ct do controle interno}) \text{ Amostra A} - (\text{Ct do gene de interesse} - \text{Ct do controle interno}) \text{ Amostra B}]}$$

Onde a Amostra A são os valores de Ct das amostras tratadas, ou seja, das plantas transgênicas; e a Amostra B são os valores de Ct das amostras não tratadas, das plantas selvagens. Utilizou-se dois pares de primers normalizadores, sendo CAC e SUI, e os cálculos

foram realizados com o par de primers com melhores valores de Ct. Também foi realizado da mesma forma para os genes de interesse, utilizando-se dois pares de primers para *SmTAT* e dois para *SvHPPR*, utilizando-se o de melhor eficiência para os cálculos de expressão gênica. Os experimentos foram realizados no laboratório de Evolução e Diversidade, campus São Bernardo do Campo da UFABC.

## **5.6 Coleta do tecido vegetal para análises bioquímicas e parâmetros biomassa e inflorescências**

Todos as linhagens obtidas, além de plantas selvagens, foram colocadas para crescimento em casa de vegetação, utilizando-se de cerca de 50 sementes maduras para cada linhagem, divididas em 10 copos, escarificadas e higienizadas com uso de hipoclorito de sódio 2%, colocadas diretamente no solo. Após 10 dias de crescimento, iniciou-se a etapa de confirmação via PCR, individualmente. A confirmação foi necessária pois nem todas as linhagens estavam na geração T2, de plantas homozigotas. A partir do resultado das PCR's foi possível dividir cada uma das linhagens em pools, contemplando todas as plantas positivas. Em geral, dividiu-se cada linhagem em 3 pools, sendo pools A, B, e C, para um total de 1 a 5 plantas representando cada pool, dependendo do número de plantas positivas. A análise de expressão gênica foi realizada posteriormente, a partir dos pools de cada linhagem, para prosseguir para os demais experimentos somente com plantas positivas, e com expressão dos genes *SmTAT* e/ou *SvHPPR*. Importante ressaltar que nem todas as plantas avaliadas estavam na geração de plantas consideradas homozigotas, por conta do tempo para avanço das plantas. Sendo assim, para E430, procedeu-se com os eventos 1, 2 e 3 na geração T2, homozigotas; e as demais em geração T1; para E432, utilizou-se o evento 1 em geração T2, e as demais em geração T1.

Para realização das análises bioquímicas de composição de parede celular, hidrólise enzimática, e tioacidólise, utilizou-se o tecido das plantas seco e macerado. Após 30-35 dias de crescimento, na fase *pré-booting*, após o crescimento vegetativo e antes da fase reprodutiva, procedeu-se para a coleta do material, separando-se folhas e colmos. Coletou-se uma porção do material para secagem em estufa, e outra porção foi prontamente congelada em nitrogênio líquido para serem posteriormente macerados também em nitrogênio para análises de metabolômica, e avaliação de AR por HPLC. O material para secagem foi acondicionado em sacos de papel, e colocado em estufa para secagem, por um período de 7 dias, a 45°C. Após a

secagem, procedeu-se para a maceração dos tecidos com auxílio de moinho de bolas modelo Retsch MM 400.

Para quantificação da biomassa seca e número de inflorescências, procedeu-se para um novo crescimento de plantas. Ponderando-se os resultados das PCR's e qPCR's realizadas anteriormente. Para sementes em geração T1, considerou-se como plantas homozigotas aquelas que apresentaram mais de 90% de plantas positivas por PCR; e sementes de geração T2 foram consideradas homozigotas, pois já haviam sido avaliadas via PCR anteriormente na etapa de avanço de geração. Na avaliação da biomassa, primeiramente procedeu-se com a pesagem do material total logo após coleta, incluindo inflorescências e sementes. Após, este mesmo material foi acondicionado em sacos de papel, e deixado para secar em estufa a 45°C por um período de 7 dias. Pesou-se novamente, e a diferença do peso inicial para o peso final foi considerado como biomassa seca total.

Os experimentos de crescimento e coleta foram realizados no laboratório de Evolução e Diversidade, campus São Bernardo do Campo da UFABC e a etapa de maceração foi realizada no LCCEM - Laboratório de Cristalografia e Caracterização Estrutural de Materiais, campus Santo André da UFABC.

## **5.7 Preparo da parede celular (CWR)**

O primeiro passo para dar início às análises de parede celular e também de hidrólise enzimática é a realização do protocolo de preparo da parede celular, descrito como *cell wall residue* (CWR). O protocolo utilizado foi desenvolvido pelo grupo de pesquisa em Bioenergia e Bioaromáticos do Instituto de Biotecnologia do Flandres (VIB), na Bélgica. Inicia-se com a pesagem das amostras utilizando-se de balança analítica e tubos eppendorf 2 mL. Pesou-se 60 ± 0,5 mg de material seco macerado (partículas de 250 a 500 µm), assim como o tubo eppendorf vazio. Incluiu-se 3 tubos como branco para cálculo da biomassa residual. Adicionou-se 1 mL de água ultrapura, homogeneizou-se em vórtex, e incubou-se em termobloco a 98°C por 30 minutos e 750 rpm. Centrifugou-se por 5 minutos a 14.000 g e 600 µL sobrenadante foram removidos. Adicionou-se 1 mL de etanol puro, homogeneizou-se em vórtex e incubou-se em termobloco a 76°C por 30 minutos a 750 rpm. Centrifugou-se por 5 minutos a 14.000 g e removeu-se 1 mL do sobrenadante. Adicionou-se 1 mL de clorofórmio, homogeneizou-se em vórtex e incubou-se em termobloco a 59°C por 30 minutos a 750 rpm. Centrifugou-se por 5 minutos a 14.000 g e removeu-se todo o sobrenadante. Adicionou-se 1 mL de acetona,

homogeneizou-se em vórtex e incubou-se em termobloco a 54°C por 30 minutos a 750 rpm. Centrifugou-se por 5 minutos a 14.000 g e removeu-se totalmente o sobrenadante. O pellet então foi seco em *speedvac* modelo RCV 2-18 CD Plus, marca CHRIST, por 4 horas a 40°C. Os experimentos foram realizados no laboratório de Evolução e Diversidade, campus São Bernardo do Campo da UFABC.

### 5.8 Lignina pelo método de Brometo de Acetila

O teor de lignina das folhas e colmos das plantas transgênicas foi determinado a partir do método do brometo de acetila, um método muito bem estabelecido na literatura. Os experimentos foram realizados no laboratório de Evolução e Diversidade, campus São Bernardo do Campo da UFABC. O protocolo utilizado foi descrito por Moreira-Vilar *et al.*, (2014). Pesou-se  $4 \pm 0,15$  mg de CWR em tubos eppendorf de 2 mL. Foram incluídos 3 tubos para representarem os brancos do experimento. Acrescentou-se 200 µL de solução brometo de acetila 25% em ácido acético glacial recém preparada e aqueceu-se em termobloco a 70°C por 30 minutos sem agitação. Aos 20 minutos, homogeneizou-se levemente os tubos. Ao fim do tempo, adicionou-se 360 µL de hidróxido de sódio 2M e resfriou-se os tubos em gelo. Adicionou-se 65 µL de cloridrato de hidroxilamina 0.5 M. Adicionou-se após 1.175 mL de ácido acético glacial e homogeneizou-se. Centrifugou-se por 5 minutos a 14.000 g. Para a leitura, utilizou-se 20 µL de amostra e 180 µL de ácido acético glacial em placa específica para leitura em UV, em absorvância de 280 nm. Para o cálculo do teor de lignina nas amostras utilizou-se a Equação 3.

#### Equação 3 – Cálculo do conteúdo de lignina método brometo de acetila

$$\text{Lignina (\%CWR)} = \frac{\text{Absorbância} \times \text{fator de diluição}}{(\text{Coeficiente de extinção} \times \text{comprimento da placa}) \times \frac{\text{Volume final da reação} \times 100}{\text{Biomassa inicial (mg CWR)}}$$

Onde:

- Absorbância: é dada pela leitura do equipamento;
- Fator de diluição: 10 (para uma diluição 20 µL de amostra e 180 µL de ácido acético);
- Coeficiente de extinção: 18,955, valor para gramíneas de acordo com Fukushima; Hatfield (2004);

- Caminho óptico: 0,5556 cm de acordo com o fabricante;
- Volume final de reação: 1,8 mL;
- Biomassa inicial:  $\pm 4$  mg

## 5.9 Lignina pelo método CASA

O método de quantificação de lignina denominado CASA vem do inglês *Cysteine-Assisted Sulphuric Acid*, e consiste em uma quantificação rápida de forma indireta, baseada em uma hidrólise com ácido sulfúrico e cisteína, dissolvendo todo o material, incluindo a lignina, em um mecanismo de reação ainda não elucidado. Os experimentos foram realizados no laboratório de Evolução e Diversidade, campus São Bernardo do Campo da UFABC. O método é descrito por Lu *et al.*, (2021). Para este método, utilizou-se de 10 mg ( $\pm 0,15$ ) de CWR, uma solução fresca de 0,1 g/mL de cisteína em ácido sulfúrico 72%, água ultrapura, frascos de 100 mL, termobloco e espectrofotômetro UV-VIS 96 poços. O primeiro passo é a pesagem da amostra, anotando-se o peso do tubo vazio, e o peso de CWR, em tubos de 2 mL. Após, adicionou-se 1 mL da solução de cisteína diluída em ácido sulfúrico, homogeneizou-se em vórtex, e incubou-se em termobloco por 1h, a 30°C, em agitação de 1.000 rpm. Após resfriar para temperatura ambiente, transferiu-se todo o material da reação para um frasco volumétrico de 100 mL, lavando-se os tubos da reação com água ultrapura por 5 vezes, e também transferindo-se a água da lavagem para os frascos. Completou-se com água de forma precisa até 100 mL. Transferiu-se 200  $\mu$ L para um poço de placa para leitura em 283 nm em espectrofotômetro UV-VIS, considerando-se triplicata para cada amostra. Para o cálculo do conteúdo de celulose utilizou-se a lei de Lambert-Beer, como é mostrado na Equação 4:

### Equação 4 – Cálculo do conteúdo de lignina método CASA

$$CASA \text{ Lignina } (\%CWR) = \frac{Abs \times V}{\epsilon \times m \times L} \times 100$$

Onde:

- Abs: Média absoluta geral das absorbâncias de uma amostra;
- V: Volume total de diluição em litros (L);

- $\epsilon$ : Coeficiente de absorção da lignina em ultravioleta. Utilizou-se  $11,23 \text{ L.g}^{-1}.\text{cm}^{-1}$  conforme autores do método indicam para materiais lignocelulósicos de monocotiledôneas;
- m: massa inicial da amostra (g);
- L: caminho ótico (cm). Para um tamanho normal de placa foi utilizado o coeficiente de  $0,5556 \text{ cm}$ .

## 5.10 Tioacidólise

A composição da lignina é bastante importante para evidenciar se houveram alterações nos teores dos monômeros G, S, ou H, e tricina das plantas transgênicas, a partir da modificação genética. Para isto, utilizou-se o método de Tioacidólise, adaptado do protocolo descrito por Chen *et al.*, (2021). Para este protocolo utiliza-se o material de CWR preparado como descrito anteriormente na seção 5.7. A partir do CWR, 2 mg de material seco e macerado foram acuradamente pesadas e colocadas em um vial de vidro. Adicionou-se 1 mL do reagente de Tioacidólise (16 mL de 1,4-Dioxano; 1,04 mL de Eterato de Boro; 4 mL de Etanetiol; 0,240 mL de 4,4'-Etilenobisfenol; e completou-se para 40 mL com 1,4-Dioxano). As amostras foram colocadas em banho maria a  $99^{\circ}\text{C}$  por um período de 4 horas, com agitações em vórtex de hora em hora. Após, as amostras foram deixadas para resfriar até atingirem temperatura ambiente. Separou-se então a solução em duas partes: uma para análise de lignina por CG-MS e outra para análise de tricina por HPLC-MS.

Para a análise de lignina por cromatografia gasosa (CG-MS), adicionou-se, em um novo vial de vidro,  $400 \mu\text{L}$  da reação anterior e misturou-se com  $190 \mu\text{L}$  de bicarbonato de sódio ( $\text{NaHCO}_3$ ) na concentração 0,4 Molar. Misturou-se a solução por pipetagem. Para secagem, os vials de vidro foram colocados em speedvac a temperatura ambiente, 700 rpm, durante toda a noite. Após, foi necessário realizar a derivatização das amostras. Para tal, adicionou-se  $100 \mu\text{L}$  do reagente de derivatização (mix de 1:1 de Piridina e BSTFA (N,O-bis(trimetilsilil)trifluoroacetamida) +1% de TMCS (trimetilclorosilano)). Fechou-se os vials e homogeneizou-se a partir de vórtex por 10 segundos. Os vials foram então colocados em banho maria a  $55^{\circ}\text{C}$  por 30 minutos. Após, homogeneizou-se novamente no vórtex, e foram deixados na banca para resfriar até atingir temperatura ambiente. Transferiu-se então  $100 \mu\text{L}$  para um novo vial com inserto para ser injetado no CG-MS.

A injeção das amostras foi realizada em sistema com equipamento CG-MS do tipo Agilent Technologies 7890A e um detector de massas seletivo do tipo MSD 5975C para análise dos monômeros de lignina. Hélio foi utilizado como gás carreador em um fluxo de 1.0 mL/min. Uma coluna do tipo HP-5MS (30 m x 0.250 mm, 0.25-micro) foi utilizada para separação dos monômeros. O volume de injeção foi mantido em 2 µL utilizando um injetor automático. A temperatura de entrada do CG foi configurada para 230°C. Realizou-se oito lavagens utilizando clorofórmio em meio à injeção das amostras. A temperatura inicial do forno foi de 100°C por 2 minutos, com uma rampa de 25°C/min até 300°C, e mantido por 3 minutos. A linha de transferência do detector de massas seletivo (MSD) foi mantido a 250°C, e a voltagem para o impacto de ionização de elétrons foi mantido em 70 eV. O escaneamento do detector de massas variou entre 50 a 500 m/z com um *delay* dBo solvente de 6 minutos e meio. Para a quantificação dos monômeros de lignina, utilizou-se os seguintes valores de tempo de retenção, como referência para cada monômero: T.R 9.85 para o padrão interno de 4,4'-etilenobisfenol; T.R 11.33 para derivados do monômero hidroxifenil (H); T.R 11.78 para derivados do monômero guaiacil (G); T.R 11.22 para derivados do monômero siringil (S). Os cálculos de quantificação foram realizados utilizando-se primeiramente a normalização dos valores de área para microgramas por miligramas de CWR, considerando-se 0,048 ug do padrão interno em 1 mL de amostra injetado no equipamento. A partir dos valores normalizados, o cálculo foi realizado a partir dos valores das áreas, considerando a biomassa inicial de CWR em miligramas.

Para a análise de tricina por cromatografia líquida (HPLC-MS), adicionou-se em um tubo eppendorf de 1,5 mL 400 µL da reação anterior e misturou-se com 190 µL de bicarbonato de sódio (NaHCO<sub>3</sub>) na concentração 0,4 Molar. Para secagem os tubos eppendorf foram colocados em termobloco a 28°C, agitação 750 rpm também durante toda a noite para secagem completa. Após secagem, adicionou-se 100 µL de uma solução metanol:dimetilsulfóxido (DMSO) 1:1, homogeneizou-se em vórtex, e as amostras foram colocadas em banho maria a 55°C por 15 minutos e agitação 700 rpm. Após, homogeneizou-se novamente em vórtex, por 10 segundos, centrifugou-se em velocidade máxima por 10 minutos e transferiu-se o sobrenadante para vial contendo inserto.

O equipamento utilizado para análise de LC-MS foi do tipo Agilent HPLC 1290 Infinity II, com um sistema contendo um injetor G729B, uma bomba G7104A, um termostato multicoluna G7116B, e um diodo detector G7117A. Uma coluna do tipo Agilent Zorbax Eclipse Plus C18 (2.1 x 50 mm, 1.8-micro) foi utilizada para separação dos compostos. A fase móvel do HPLC foi composta por uma solução de água com 0,1% de ácido fórmico (v/v) como

solvente A, e uma solução de acetonitrila com 0,1% de ácido fórmico (v/v) como solvente B. O termostato da coluna foi mantido em 30°C com um fluxo de solvente de 0,45 mL/min. O gradiente de eluição foi de 6% de B por 3 minutos; de 6% a 95% de B em 5 minutos; mantido em 95% de B por 2 minutos, e após, de 95% para 6% de B em 1 minuto e mantido em 6% de B por 2 minutos. O tempo total de corrida foi de 13 minutos por amostra. Para identificação, foi utilizado um espectrômetro de massas do tipo triplo quadrupolo modelo Agilent 6460C, configurado para o escaneamento e quantificação de tricina. Gás de nitrogênio foi usado em um fluxo de 11 L/min a 350°C. A temperatura do gás de capilaridade foi configurada para 300°C, a um fluxo de 10 L/min, e pressão 45 psi. A voltagem da ionização por eletrospray foi mantida em 3,500 v no modo de ionização positiva e 3.500 v no modo de ionização negativa. A voltagem do fragmentador foi mantida em 135 v. Um padrão de tricina de concentração 0,04 mg/mL foi utilizado, e os cálculos feitos a partir da sua curva padrão, observando-se os valores de Mau no comprimento de onda em 352nm, e tempo de retenção 16,67 min, que são os respectivos tempos de retenção e comprimento de onda do padrão de tricina.

As etapas de extração foram realizadas no laboratório de Evolução e Diversidade, campus São Bernardo do Campo da UFABC. As demais etapas dos protocolos, como a utilização do CG-MS e do HPLC-MS foram realizadas no Departamento de Botânica - IB/USP, em São Paulo.

### 5.11 (Hemi)Celulose

Para determinação do conteúdo de hemicelulose e celulose, utilizou-se o protocolo descrito por Foster; Martin; Pauly (2010). Os experimentos foram realizados no laboratório de Evolução e Diversidade, campus São Bernardo do Campo da UFABC. Iniciou-se com a remoção da celulose por meio do reagente TFA – Ácido trifluoroacético 2 M: 30,95 mL de TFA e 169,05 mL de água ultrapura autoclavada. Primeiramente, pesou-se tubos vazios de 2 mL e  $10 \pm 0,15$  mg de CWR previamente preparado, e incluiu-se 3 tubos vazios para o branco do experimento. Na capela, adicionou-se 1 mL de TFA 2 M, e incubou-se em termobloco a 99°C por 2 horas e 750 rpm. Centrifugou-se a 18.000 g por 10 minutos. Descartou-se o sobrenadante e procedeu-se para a lavagem do material primeiramente com água ultrapura autoclavada, por duas vezes, e após, com acetona, também por duas vezes, fazendo-se uso de vórtex, e centrifugação a 18.000 g por 10 minutos. Após a lavagem, secou-se o pellet em *speedvac* overnight a 40°C e pesou-se novamente os tubos para cálculo do conteúdo de hemicelulose,

utilizando-se da biomassa inicial e biomassa final. Para determinação do conteúdo de celulose, utilizou-se o reagente Updegraff (ácido acético:ácido nítrico:água; 8:1:2 v/v). Inicialmente, incluiu-se tubos vazios para o banco e adicionou-se 1 mL do reagente ao pellet obtido na etapa de hemicelulose. Procedeu-se para o vórtex e as amostras foram colocadas em termobloco, dentro da capela, a 99°C por 30 minutos, sem agitação. Somente a celulose cristalina permanece insolúvel no pellet após esta etapa. Após, resfriou-se as amostras em gelo até atingirem temperatura ambiente. Centrifugou-se a 18.000 g por 15 minutos e descartou-se 900 µL de sobrenadante. Adicionou-se 1 mL de água, centrifugou-se a 18.000 g por 15 minutos e descartou-se 1 µL do sobrenadante. Realizou-se três lavagens com acetona pura, seguidas de centrifugação a 18.000 g por 15 minutos e remoção do sobrenadante. Levou-se então os tubos para o *speedvac* por 30 minutos a 40°C.

A próxima etapa se trata da hidrólise de Saeman, onde adicionou-se 350 µL de ácido sulfúrico 72% ao pellet e incubou-se a temperatura ambiente por 30 minutos, sem agitação. Procedeu-se para a homogeneização em vórtex e os tubos foram deixados para mais 15 minutos de incubação. Adicionou-se então 1,65 mL de água ultrapura autoclavada, homogeneizou-se em vórtex e centrifugou-se a 18.000 g por 10 minutos. Coletou-se 700 µL do sobrenadante em novos tubos e procedeu-se para a quantificação de celulose a partir do método da Antrona, recém preparada (2 mg de Antrona por 1 mL de ácido sulfúrico).

Para a leitura, utilizou-se da reação em tubos, onde adicionou-se 20 µL de sobrenadante e 480 µL de água ultrapura autoclavada. Colocou-se os tubos no gelo e após, adicionou-se 1 mL de Antrona recém preparada. Homogeneizou-se em vórtex e incubou-se em termobloco por 30 minutos a 80°C. Transferiu-se 200 µL da reação para placa de leitura de 96 poços e a absorbância foi lida em 625 nm. O cálculo de glicose foi realizado a partir de uma curva padrão previamente estabelecida utilizando-se de diluições seriadas de glicose.

## **5.12 Pré-tratamento de biomassa e sacarificação**

Para o pré-tratamento de biomassa, utilizou-se o protocolo descrito por Acker *et al.*, (2016), utilizando o pré-tratamento alcalino com hidróxido de sódio (NaOH) 50 mM. Os experimentos foram realizados no laboratório de Evolução e Diversidade, campus São Bernardo do Campo da UFABC. Pesou-se tubos de 2 mL vazios e  $10 \pm 0,5$  mg de biomassa seca, e adicionou-se 1 mL de NaOH 50 mM. Incubou-se em termobloco a 90°C por 3 horas e agitação de 750 rpm. Resfriou-se os tubos em gelo por 5 minutos, e centrifugou-se a 35541 g

por 5 minutos. Removeu-se o sobrenadante e o pellet foi lavado duas vezes utilizando-se água ultrapura autoclavada, seguido de vórtex e centrifugação a 35541 g por 5 minutos. Após a remoção do sobrenadante, adicionou-se 1 mL de etanol puro e incubou-se em termobloco a 80°C por 30 minutos e agitação de 750 rpm. Centrifugou-se a 35541 g por 5 minutos e removeu-se o sobrenadante. Adicionou-se novamente etanol puro e procedeu-se para uma nova lavagem. Duas etapas de lavagem com etanol foram realizadas. Após a remoção do etanol, procedeu-se para uma nova lavagem com acetona, adicionando-se 1 mL de acetona pura e incubação em termobloco a 54°C por 30 minutos e agitação de 750 rpm. Após centrifugação e remoção do sobrenadante, o pellet foi colocado em *speedvac* por um período de 4 horas a 40°C. Os tubos foram então pesados após o pré-tratamento para o cálculo de recuperação de biomassa. Para amostras sem pré-tratamento, realizou-se somente a etapa de lavagem, como descrita anteriormente, a partir de lavagens com etanol e acetona.

Para a sacarificação utilizou-se o protocolo descrito por Oliveira *et al.*, (2020) utilizando-se amostras provenientes do protocolo de pré-tratamento de biomassa. Primeiramente, pesou-se 10 mg  $\pm$  0,15 mg de material em tubos 2 mL. Preparou-se o tampão acetato 50 mM pH 4.8 e 0,02% (v/v) de azida de sódio. Adicionou-se 950  $\mu$ L de tampão acetato 50 mM nas amostras com e sem pré-tratamento, e incubou-se em termobloco a 50°C e agitação de 750 rpm. Após cinco minutos, adicionou-se 50  $\mu$ L do coquetel enzimático de celulase Cellic CTec2, marca Novozymes, em uma diluição de 100x em tampão acetato, com atividade previamente determinada. Em diversos pontos, determinados tendo em vista o objetivo de cada experimento, foram realizadas as leituras. Após colocar os tubos no gelo, recolheu-se 20  $\mu$ L de sobrenadante em uma placa plástica de PCR 96 poços, contendo 30  $\mu$ L de água ultrapura autoclavada. Adicionou-se após 50  $\mu$ L de DNS (Ácido 3,5-dinitrosalicílico) previamente preparado, e incubou-se em termociclador por 5 minutos a 99°C. Após resfriamento até a temperatura ambiente, adicionou-se 100  $\mu$ L de água ultrapura autoclavada na placa de PCR, e ainda outros 100  $\mu$ L de água ultrapura autoclavada na placa de leitura. Retirou-se 100  $\mu$ L da mistura da placa de PCR e adicionou-se à água da placa de leitura, totalizando 200  $\mu$ L. A leitura foi realizada em espectrofotômetro em comprimento de onda de 525 nm, e a concentração de açúcares foi medida a partir de uma curva de glicose previamente estabelecida.

Para alguns casos, também se utilizou de outros métodos para a quantificação de açúcar, com a finalidade de confirmar o resultado obtido a partir do DNS. O método de DNS é interessante na quantificação de açúcares redutores, e um outro método para quantificação de açúcares totais, como o método Glicose Liquiform, pode ser interessante na confirmação do

resultado da sacarificação. Primeiramente, preparou-se uma curva padrão com uma solução de glicose variando de 0,25 mg/mL até 1,0 mg/mL. Para quantificação das amostras utilizou-se o sobrenadante das amostras de sacarificação após 24 horas de reação, e a equação da reta da curva gerada. Centrifugou-se as amostras por 10 minutos, temperatura ambiente em velocidade máxima. Conforme indicado no protocolo de utilização, adicionou-se 1 mL do Reagente 1, que é um coquetel enzimático, a 0,01 mL do sobrenadante das amostras de sacarificação. Misturou-se vigorosamente com auxílio de vórtex, e as amostras foram então incubadas a 37°C em banho-seco (termobloco) por 10 minutos. Após, procedeu-se para a leitura em espectrofotômetro de cubeta no comprimento de onda de 505 nm.

Além disso, o método do reagente de Benedict descrito por Hernández-Lopes *et al.* (2020) também foi utilizado para confirmação dos resultados da sacarificação, na quantificação de açúcares redutores totais. Preparou-se uma curva padrão de glicose variando de 0,25 mg/mL até 1,0 mg/mL. Para o reagente de Benedict, utilizou-se citrato de sódio, carbonato de sódio e sulfato de cobre. A reação foi composta por 100 µL de reagente para 50 µL de sobrenadante de sacarificação, em placa de PCR. Utilizou-se de termociclador a 99°C por 5 minutos. Após o resfriamento da placa, transferiu-se 100 µL para uma placa de leitura em espectrofotômetro em comprimento de onda de 740 nm, os resultados então foram obtidos a partir da equação da reta da curva padrão de glicose.

### **5.13 Fenólicos Totais**

Para a quantificação da presença de fenólicos totais nas amostras de tecido vegetal das plantas, utilizou-se o método descrito por Silveira *et al.*; (2018), em um documento da Embrapa, de forma adaptada, para reduzir a quantidade de biomassa utilizada. O método é baseado em uma reação de oxidação-redução em condições alcalinas, em que o íon fenolato é oxidado, enquanto o reagente de Folin é reduzido. Para tal, primeiramente preparou-se uma solução de metanol 50% (v/v), uma solução de acetona 70% (v/v), e foi realizada uma diluição do reagente de Folin Ciocalteu comercial em uma proporção 1:3, todos em água ultrapura. Preparou-se ainda uma solução de carbonato de sódio anidro p.a 20%, onde utilizou-se 20 g de carbonato de sódio dissolvidos em 100 mL de água ultrapura. Além disso, preparou-se uma solução de ácido gálico p.a, sendo 5 mg de ácido gálico monoidratado p.a dissolvidos em 100 mL de água ultrapura.

Procedeu-se para o preparo da curva padrão de ácido gálico, a partir de diluições seriadas para os pontos 0, 10, 20, 30, 40 e 50  $\mu\text{g}$  de massa de ácido gálico. Partiu-se então para a reação, em triplicata, adicionando-se a cada tubo Falcon de 15 mL, 1 mL de reagente Folin (1:3); 2 mL da solução de carbonato de sódio 20%; e 2 mL de água ultrapura. Homogeneizou-se em vórtex e os tubos foram colocados em repouso por 30 minutos em temperatura ambiente, em local com baixa luminosidade. Para a leitura dos pontos, transferiu-se alíquotas de 200  $\mu\text{L}$  para uma placa de leitura em espectrofotômetro a 700 nm, utilizando como branco a solução sem a presença de ácido gálico (0,0  $\mu\text{g}$ ). Procedeu-se para a leitura, e obtenção da curva com a equação da reta que foi utilizada para quantificação dos fenólicos a partir da reação com ácido gálico das amostras de tecidos vegetais.

Para a obtenção do extrato utilizado na reação procedeu-se primeiramente com uma extração metanólica, utilizando-se 100 mg de material, onde foi adicionado 400  $\mu\text{L}$  de metanol 50%, que após homogeneização foram colocados em repouso no escuro por 1h em temperatura ambiente. Após, centrifugou-se em velocidade máxima por 15 minutos, filtrou-se e transferiu-se para um novo tubo. Ao resíduo de extração foram adicionados 400  $\mu\text{L}$  de acetona 70%, homogeneizou-se em vórtex, e deixou-se em repouso por 1h em temperatura ambiente. Centrifugou-se da mesma forma, filtrou-se e o filtrado foi combinado com o filtrado anterior. Para reação utilizando as amostras dos extratos vegetais procedeu-se da mesma forma que para a realização da curva padrão. O cálculo foi realizado levando-se em conta a equação da reta, os valores de absorbância e o peso inicial das amostras em miligramas. Os experimentos foram realizados no laboratório de Evolução e Diversidade, campus São Bernardo do Campo da UFABC.

### **5.14 Ácido Rosmarínico por HPLC**

Para verificar a presença ou não de ácido rosmarínico nas plantas transgênicas e nas plantas selvagens, procedeu-se para uma avaliação dos extratos vegetais a partir de HPLC. Para a obtenção dos extratos foram utilizadas de 100 a 150 mg de biomassa fresca de colmos, mantidos a  $-80^{\circ}\text{C}$  após a coleta. Após pesar o material, adicionou-se 1 mL de uma solução de metanol 70%, homogeneizou-se em vórtex por 2 minutos, e depois incubou-se as amostras em termobloco a  $40^{\circ}\text{C}$  por 1h, com agitação suave a 300 rpm. Após, centrifugou-se em velocidade máxima por 10 minutos e removeu-se 1 mL do sobrenadante, que foi armazenado a  $-20^{\circ}\text{C}$ .

Para a leitura do extrato em HPLC UV-Vis, utilizou-se um equipamento da marca AKTA Purifier 900, e uma coluna de fase reversa do tipo VYDAC C18 de dimensões 250 mm × 4,6 mm, 5 µm. Utilizou-se um gradiente de tampão A (0,1% Ácido Acético em água ultrapura) e tampão B (Acetonitrila), sendo 0–10 min: 20-40% B; 10–20 min: 40-60% B; 20–25 min: 60-20% B, em um fluxo de 1 mL/min e temperatura de 25°C. O volume de injeção foi de 1 mL e o detector UV-Vis foi configurado para um comprimento de onda de 330 nm, sugerido na literatura para detecção de AR (LIU et al., 2013). Os testes foram realizados no Laboratório de Bioquímica e Biologia Molecular do campus São Bernardo do Campo da UFABC.

Para a quantificação de AR utilizou-se uma curva da calibração a partir de um padrão comercial de AR (Sigma Aldrich), nas diluições de 0,5; 2; 10; 50 µg/mL. Usando a equação da reta da curva de calibração, os valores de mAU (mil unidades de absorbância) foram definidos a partir da área correspondente do pico do padrão de AR. Usando os tempos de retenção da área, fez-se o cálculo da média de mAU, cujo valor foi utilizado na equação da reta e normalizado a partir do peso inicial da amostra. Os resultados foram expressos então em mg/g de biomassa inicial.

## 5.15 Metabolômica

Para a avaliação do perfil de metabólitos das linhagens transgênicas das duas construções, todas as linhagens foram avaliadas, mas os resultados foram discutidos levando-se em conta as linhagens com resultados mais significativos nos experimentos anteriores, como lignina e sacarificação. Neste caso também se avaliou os colmos das plantas, a partir de uma extração metanólica utilizando-se de tecido fresco, que foi mantido em -80°C após a coleta. O protocolo foi definido pelo grupo do Laboratório Nacional de Biorenováveis (LNBR) do CNPEM - Centro Nacional de Pesquisa em Energia e Materiais, que também realizou os experimentos.

Iniciou-se com a pesagem de tubos eppendorf 2 mL vazios, e de 50 a 200 mg de biomassa fresca em cada tubo. Adicionou-se 1,5 mL da solução extratora de metanol 80% (v/v), e agitou-se em vórtex em velocidade máxima por 15 minutos. Após a agitação, centrifugou-se por 15 min a 20.000 g e 4°C, transferiu-se o sobrenadante para um novo tubo, e centrifugou-se novamente nas mesmas condições. O sobrenadante foi transferido para um terceiro tubo previamente pesado, e os tubos foram colocados em speedvac em temperatura ambiente até

completa secagem do extrato. Após a secagem pesou-se novamente os tubos para registrar a massa do extrato de cada amostra.

Na próxima etapa, procedeu-se para a extração em fase sólida utilizando cartuchos de SPE Oasis (Waters). As amostras foram reconstituídas com uma solução de equilíbrio/lavagem composta por Metanol 5 % (v/v) e Ácido Fórmico 1 % (v/v). Condicionou-se os cartuchos utilizando-se de 1 mL de metanol 100%, após, adicionou-se 1 mL da solução de equilíbrio/lavagem, o procedimento foi repetido por 2 vezes. Partiu-se para o carregamento da amostra, adicionando-se 1 mL do extrato na coluna SPE, mantendo-se a vazão baixa em 1 mL/min e repetindo-se por 2 vezes. Procedeu-se para a lavagem do cartucho com 1 mL da solução de equilíbrio lavagem. Após, na etapa de eluição, adicionou-se 1 mL de metanol 100% com vazão baixa e coletou-se o material em um novo tubo. Procedeu-se para a secagem das amostras com uso de speedvac em temperatura ambiente, e depois as amostras foram congeladas e enviadas para análise.

Para a análise dos metabólitos utilizou-se de LC-MS/MS nos modos positivo e negativo. A solubilização das amostras foi feita em solução de acetonitrila 5%, adicionando-se o volume necessário para obter a concentração de 1 mg/mL. As amostras foram agitadas em agitador vórtex por 10 minutos e centrifugadas por 10 minutos a 6720 g. O sobrenadante foi coletado e filtrado em microfiltro de PTFE 0,22 µm. O cromatógrafo utilizado foi do tipo Waters UPLC ACQUITY Premier, com fase estacionária: Waters ACQUITY UPLC HSS T3 100 mm x 2,1 mm, 1,8 µm; e Fase móvel composta por A: ácido fórmico 0,1 % v/v, B: acetonitrila + ácido fórmico 0,1 % v/v. O gradiente utilizado foi de: 0-3 min: 95% A e 5%B; 15min: 50% A e 50% B; 18-20 min: 100% B; 21-23 min: 95% A e 5% B. A vazão foi mantida em 0,280 mL/min, o tempo de análise de 23 min para um volume de injeção de 5 µL, temperatura do forno a 40 °C e temperatura do amostrador a 6 °C.

O espectrômetro de massas utilizado foi do tipo Waters Synapt XS, acoplado a analisadores qToF (quadrupolo e tempo de voo) para um modo de ionização por Eletrospray em modo positivo e negativo. A faixa de varredura vaiou de m/z 50 a 1200, em um volume do capilar de 2,50 kV, voltagem de cone de 70 V, temperatura da fonte de 120°C, temperatura de dessolvatação de 500°C, vazão de gás N<sub>2</sub> do cone de 100 L/h, e do gás N<sub>2</sub> de dessolvatação de 1000 L/h, para uma pressão de gás do nebulizador de 5,5 bar e energia de colisão com uma rampa de 15 a 50 eV. O modo de aquisição dos dados foi o Data-Dependent Acquisition (DDA): seleção dos cinco íons mais intensos de cada espectro para fragmentação (MS/MS).

Para o processamento de dados utilizou-se o software Waters Progenesis QI, e adotou-se como critério de seleção de picos aqueles com intensidade maior ou igual a 1% do pico base do espectro, utilizando a base de dados Waters METLIN MS/MS Library 2019. O parâmetro utilizado foi a busca de moléculas contendo os elementos C, H, O, N, S, P, Na e K com tolerância de 10 ppm para íon precursor e fragmento. Para a anotação da possível identificação dos compostos utilizou-se os valores de m/z anotados na literatura no modo negativo de compostos fenólicos e envolvidos na biossíntese de lignina e AR.

### 5.16 Tratamento estatístico

Os gráficos foram gerados a partir do software GraphPad Prism 8. As análises estatísticas foram realizadas dentro do software. Para as análises bioquímicas, foram utilizados pools para cada um dos eventos avaliados, cada um deles representado por um número de plantas de um mesmo evento. Utilizou-se o tratamento estatístico baseado em distribuição Gaussiana em testes paramétricos, utilizando teste T não pareado, com correção de Welch e intervalo de confiança de 95% ( $P < 0.05$ ). Nos gráficos, a ilustração da significância é dada como \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ; \*\*\*\*,  $P < 0.0001$ . Os valores de erro apresentados se referem ao erro padrão a partir da média (SEM).

## 6. Resultados

### 6.1 Similaridade dos genes da via de biossíntese de AR em *S. viridis*

A presença da maquinaria de biossíntese de AR em *S. viridis* ainda não foi descrita na literatura. Sendo assim, a incerteza sobre a presença e a produção de ácido rosmarínico endógeno pela espécie permanecem. Para elucidar inicialmente esta questão, foi realizada a busca dos genes similares que compõem a via de AR em *S. miltiorrhiza* na espécie *S. viridis*. Para a busca dos genes, utilizou-se a sequência de aminoácidos dos genes de referência indicados na literatura. Como pode ser visto na Tabela 1 a seguir, grande parte da maquinaria de biossíntese de AR está presente em *S. viridis*, variando-se os valores de identidade e e-value para cada uma das enzimas chave da rota. Para o gene *PAL* (Phenylalanine ammonia-lyase) em *S. viridis* obteve-se um retorno de 11 hits diferentes, todos com e-values de 0, variando em identidades de 68 a 76%. Para o gene *C4H* (Cinnamate-4-hydroxylase), a busca pela sequência

de aminoácidos resultou em 100 hits, sendo 3 deles com *e-values* de 0 e identidade de 64 a 86%. Para *4CL* (4-Coumarate:CoA ligase), a busca retornou 80 hits, sendo 6 com *e-values* de 0, com identidade variando entre 60 e 67%. Já para *RAS* (Rosmarinic acid synthase), a busca evidenciou 100 hits, com o maior valor de identidade de 49% e *e-value* de 1.62e-142. Para *TAT* (Tyrosine aminotransferase) a busca revelou 49 hits, sendo o resultado com maior identidade de 66% e um *e-value* de 1.79e-23. Para *HPPR* (Hydroxyphenylpyruvate reductase), encontrou-se 19 hits, sendo o gene utilizado neste trabalho o resultado com maior identidade, de 70%, e um *e-value* de 4.88e-147. Para *CYP98A* (Cytochrome P450 hydroxylase family) houve o retorno de 100 hits, mas somente um com *e-value* de 0, sendo este com identidade de 69%.

**Tabela 1** – Enzimas chave na biossíntese de Ácido Rosmarínico.

Enzimas chave da biossíntese de AR	ID da proteína QUERY em <i>S. multiorrhiza</i>	Hits	ID mais similar em <i>S. viridis 4.1</i> pelo BLAST	Identidade (%)	Positivos (%)	E-value	Referência do identificador do gene inicial
<b>PAL (Phenylalanine ammonia-lyase)</b>	GenBank: EF462460.1	11	Phytozome: Sevir.1G245232.1.p	76	85%	0	Hou <i>et al.</i> , (2013)
<b>C4H (Cinnamate-4-hydroxylase)</b>	GenBank: ABC75596.1	100	Phytozome: Sevir.5G366500.1.p	86	93%	0	Zhang <i>et al.</i> , (2020)
<b>4CL (4-Coumarate:CoA ligase)</b>	GenBank: AAP68991.1	80	Phytozome: Sevir.4G201300.1.p	67	83%	0	Zhang <i>et al.</i> , (2020)
<b>TAT (Tyrosine aminotransferase)</b>	GenBank: ABC60050.1	49	Phytozome: Sevir.1G135300.1.p	66	85%	1.79e-23	Xiao <i>et al.</i> , (2011)
<b>RAS (Rosmarinic acid synthase)</b>	GenBank: ADA60182.1	100	Phytozome: Sevir.7G164100.1.p	49	66%	1.62e-142	Zhang <i>et al.</i> , (2020)
<b>HPPR (Hydroxyphenylpyruvate reductase)</b>	GenBank: AAZ67354.1	19	Phytozome: Sevir.5G009200.2.p	70	84%	4.88e-147	Xiao <i>et al.</i> , (2011)
<b>CYP98A (Cytochrome P450 hydroxylase family)</b>	GenBank: ADP00279.1	100	Phytozome: Sevir.3G199000.1.p	69	83%	0	Zhang <i>et al.</i> , (2020)

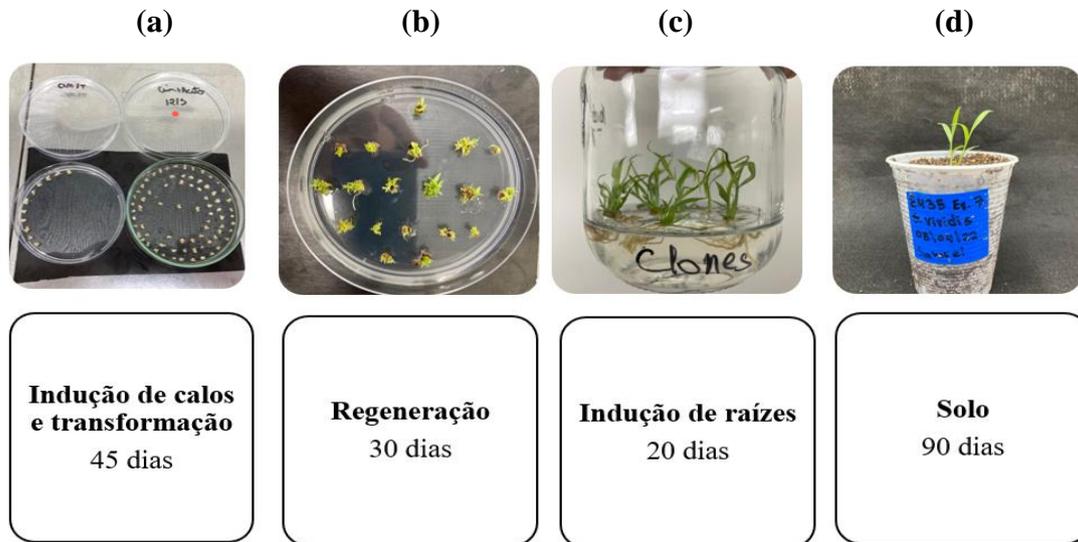
**Fonte:** elaboração própria (2025).

## 6.2 Transformação genética e caracterização molecular

A Figura 8, abaixo, e a Tabela 9, no Anexo I, evidenciam as etapas de transformação, descrevendo a quantidade de dias necessários para cada uma das etapas, assim como todas as rodadas de transformação realizadas, respectivamente. Ao todo, após 43 transformações realizadas, foi possível obter 8 eventos transformados da construção E430

*pZmUbi::Smtat::p35Sm::Svhppr;* e outros 10 eventos da construção E432  
*pIRX5::Smtat::pIRX1::Svhppr.*

**Figura 8** - Etapas do processo de transformação genética de *S. viridis*.

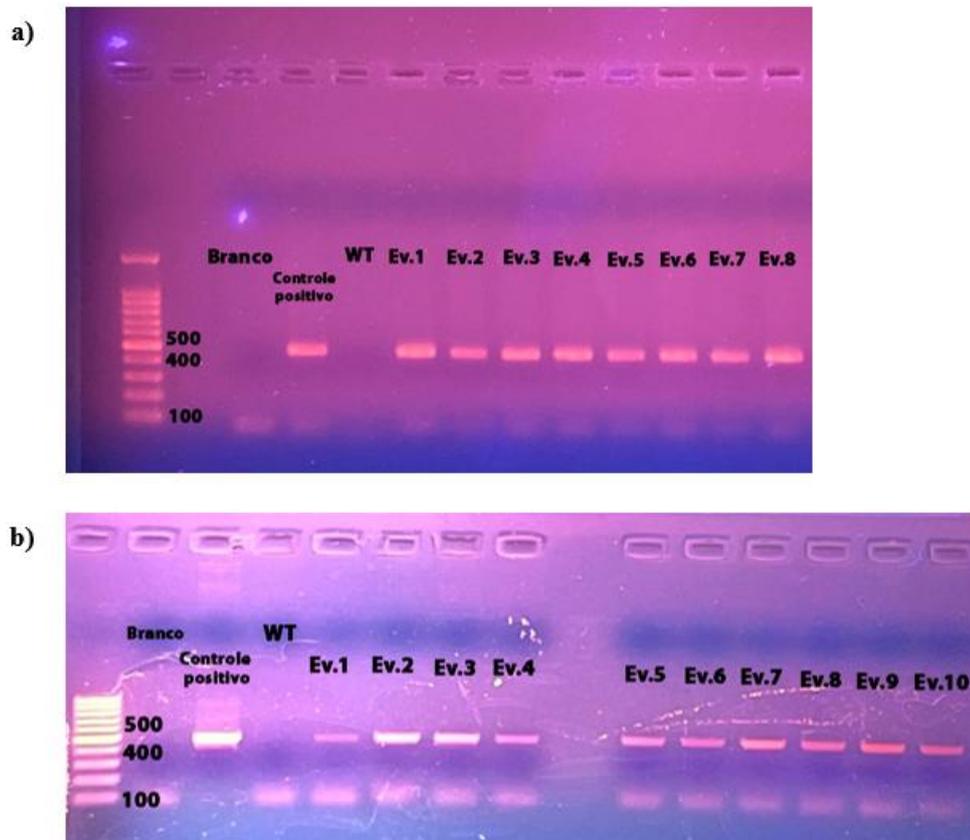


**Nota:** (a) Indução de calos após escarificação e higienização de sementes; (b) regeneração na luz após trocas de meio com antibióticos de seleção; (c) indução de raízes após regeneração da parte aérea; (d) transferência do explante para o solo e casa de vegetação.

**Fonte:** do autor (2024).

Para confirmação dos eventos T0 gerados de ambas as construções, utilizou-se de uma reação de PCR como descrito anteriormente, com o uso de primers *pAct*, cujo *amplicon* é de 481 pares de base. A Figura 9 demonstra que as plantas positivas para ambas as construções amplificaram na altura correta de banda. Estas plantas foram denominadas T0, e aguardaram até a senescência em casa de vegetação, para coleta de sementes e posterior avanço de geração.

**Figura 9** – Gel de eletroforese da reação de PCR com primers *pAct* dos eventos transformados de E430 e E432.



**Nota:** (a) Linhagens de E430; (b) Linhagens de E432. O branco utilizado refere-se ao mix de reação com água substituindo o *template* de DNA. O controle positivo refere-se à uma amostra de miniprep de cada construção.

**Fonte:** do autor (2024).

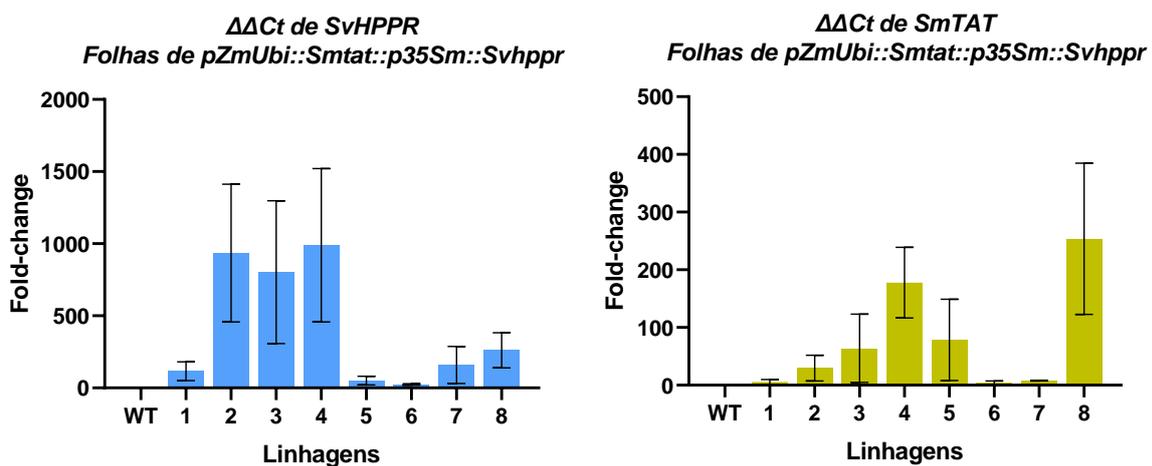
Conforme descrito anteriormente, também se utilizou de PCR's para determinar as plantas positivas em geração T1, para a fenotipagem e avaliação da expressão gênica. As plantas avaliadas, os resultados das PCR's e a divisão dos pools são demonstrados nas Tabelas 10, 11, 12 e 13 no anexo I.

Para analisar a expressão gênica e verificar o aumento da expressão dos genes de interesse da construção E430, foram realizados os protocolos de extração e tratamento de RNA, síntese de cDNA e qPCR. Os resultados das etapas prévias à qPCR não são apresentados devido ao grande volume de amostras. O experimento avaliou a expressão dos genes *SvHPPR* e *SmTAT* em plantas transgênicas comparadas às selvagens. Os primers foram escolhidos com base na literatura (Martins *et al.*, 2016). A análise considerou folhas, assumindo correlação com colmos,

devido aos promotores constitutivos da construção E430 *pZmUbi::Smtat::p35Sm::Svhppr*. A avaliação foi feita a partir dos pools de cada linhagem.

Os resultados do cálculo para a expressão gênica dos eventos da construção E430, para *SvHPPR* e *SmTAT* são evidenciados na Figura 10. São resultados obtidos a partir do cálculo com o gene endógeno *CAC*, utilizando a média dos valores. Como pode ser visto, todas as linhagens avaliadas demonstraram algum grau de expressão do gene *SvHPPR*, mas com bastante variação entre as linhagens. Para o gene *SmTAT*, verificou-se um comportamento similar, mas com valores de *fold-change* menores que aqueles vistos para *SvHPPR*, sendo 5,6x para a linhagem 1; 4,3x para a linhagem 6; e 7,9x para a linhagem 7.

**Figura 10** – Expressão gênica de *SvHPPR* e *SmTAT* para folhas de linhagens E430.



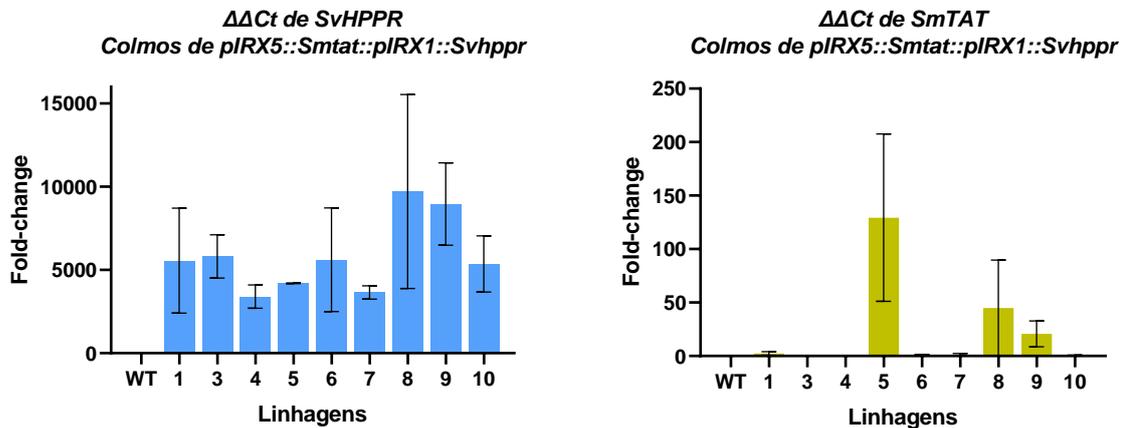
**Nota:** Cada linhagem é representada por pools, sendo pools A, B e C, totalizando um n = 3. Média dos valores calculados de  $\Delta\Delta Ct$  de cada um dos pools, para cada linhagem. Para WT considerou-se como zero. Barras de erro se referem ao erro padrão a partir da média (SEM).

**Fonte:** do autor (2024).

Para linhagens da construção E432 também foram avaliadas a expressão gênica dos genes *SvHPPR* e *SmTAT* de todas as linhagens, a partir dos pools de cada uma delas. A expressão gênica de *SvHPPR* evidenciou valores de *fold-change* expressivos para todas as linhagens avaliadas, mesmo com uma variação entre as amostras de uma mesma linhagem. Para o gene *SmTAT* somente as linhagens 1, 5, 6, 7, 8 e 9 demonstraram aumento de expressão, com destaque para as linhagens 5, 8 e 9. As linhagens 1, 6 e 7 demonstraram baixo *fold-change* em comparação com as demais linhagens, com valores menores a 2x, conforme pode ser visto na

Figura 11 a seguir. Todas as linhagens com expressão de um ou dois dos genes foram selecionadas para os experimentos de sacarificação.

**Figura 11** - Expressão gênica *SvHPPR* para colmos de linhagens E432.



**Nota:** Cada linhagem é representada por pools, sendo pools A, B e C, totalizando um n = 3. Média dos valores calculados de  $\Delta\Delta\text{Ct}$  de cada um dos pools, para cada linhagem. Para WT considerou-se como zero. Barras de erro se referem ao erro padrão a partir da média (SEM).

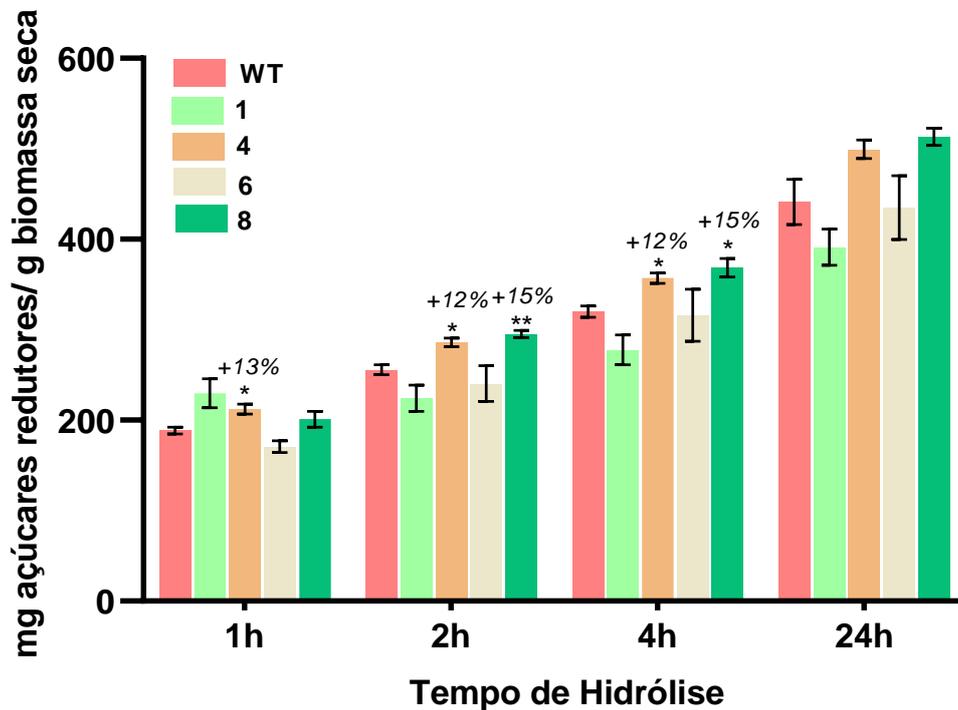
**Fonte:** do autor (2024).

### 6.3 Pré-tratamento e Sacarificação de Biomassa

Para a avaliação da hidrólise enzimática utilizou-se duas condições diferentes para as amostras, sem pré-tratamento e com pré-tratamento alcalino com NaOH 50 mM, avaliando-se a biomassa bruta e o CWR. Foram testados os colmos de todas as linhagens de E430 e os colmos de E432. Este experimento foi realizado como um *screening* para definir quais as linhagens seriam avaliadas nos experimentos posteriores.

Como pode ser visto na Figura 12, a partir da avaliação de todas as linhagens de E430, utilizando-se a biomassa bruta sem pré-tratamento, quatro linhagens apresentaram maior liberação de açúcar, sendo resultados estatisticamente significativos em comparação com as plantas selvagens somente para as linhagens 4 e 8, nos tempos 2h e 4h.

**Figura 12** – Hidrólise enzimática da biomassa bruta de colmos de linhagens E430 em condição sem pré-tratamento.

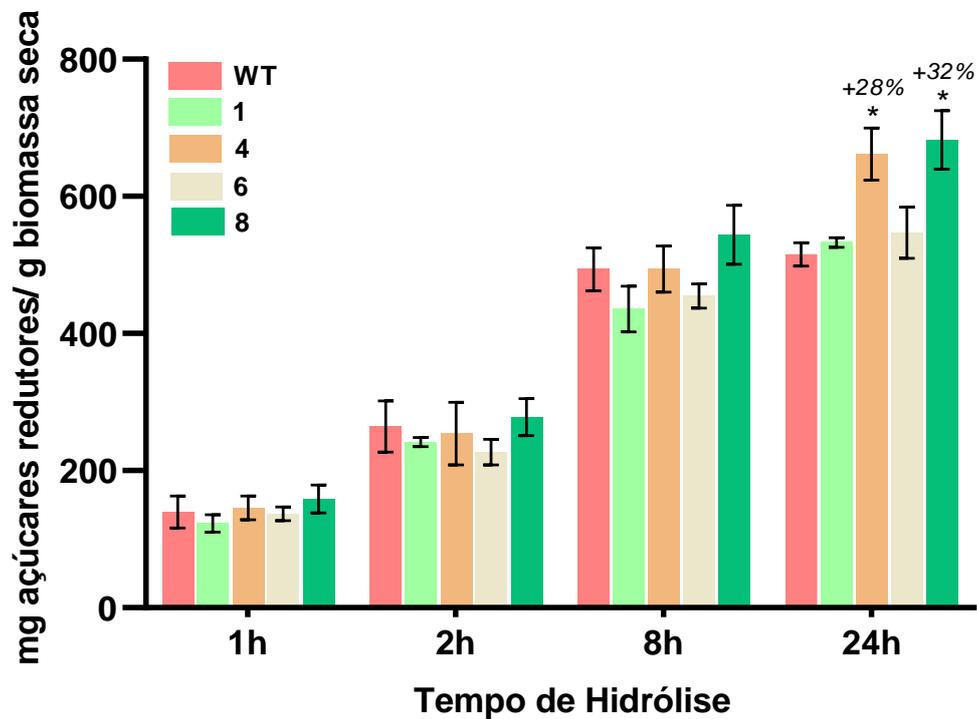


**Nota:** Cada linhagem é representada por pools, sendo pools A, B e C, totalizando um n = 3. Valores acima das barras indicando porcentagem de liberação de açúcar em comparação com WT para cada tempo de hidrólise. Barras de erro se referem ao erro padrão a partir da média (SEM). Asteriscos indicando significância estatística, sendo \*, P < 0.05; \*\*, P < 0.01; \*\*\*, P < 0.001; \*\*\*\*, P < 0.0001.

**Fonte:** do autor (2024).

Para a confirmação deste resultado, procedeu-se para a avaliação do material de CWR daquelas linhagens cujo resultados anteriores demonstraram maior liberação de açúcar sendo estatisticamente significativos, neste caso linhagens 4 e 8, e também se incluiu outras duas linhagens, 1 e 6. Como pode ser visto na Figura 13, no experimento utilizando CWR, sem pré-tratamento, as linhagens 4 e 8 demonstraram uma maior liberação de açúcar no tempo de hidrólise de 24h, sendo resultados estatisticamente significativos, confirmando os resultados com biomassa bruta. Na avaliação das linhagens 1, 4, 6 e 8 na condição de pré-tratamento alcalino para CWR, não foi observada nenhuma diferença estatisticamente significativa de liberação de açúcar em comparação com as plantas selvagens para nenhuma das linhagens avaliadas, como pode ser visto no Anexo IV – Figura 1.

**Figura 13** – Hidrólise enzimática de CWR de colmos de linhagens E430 em condição sem pré-tratamento.



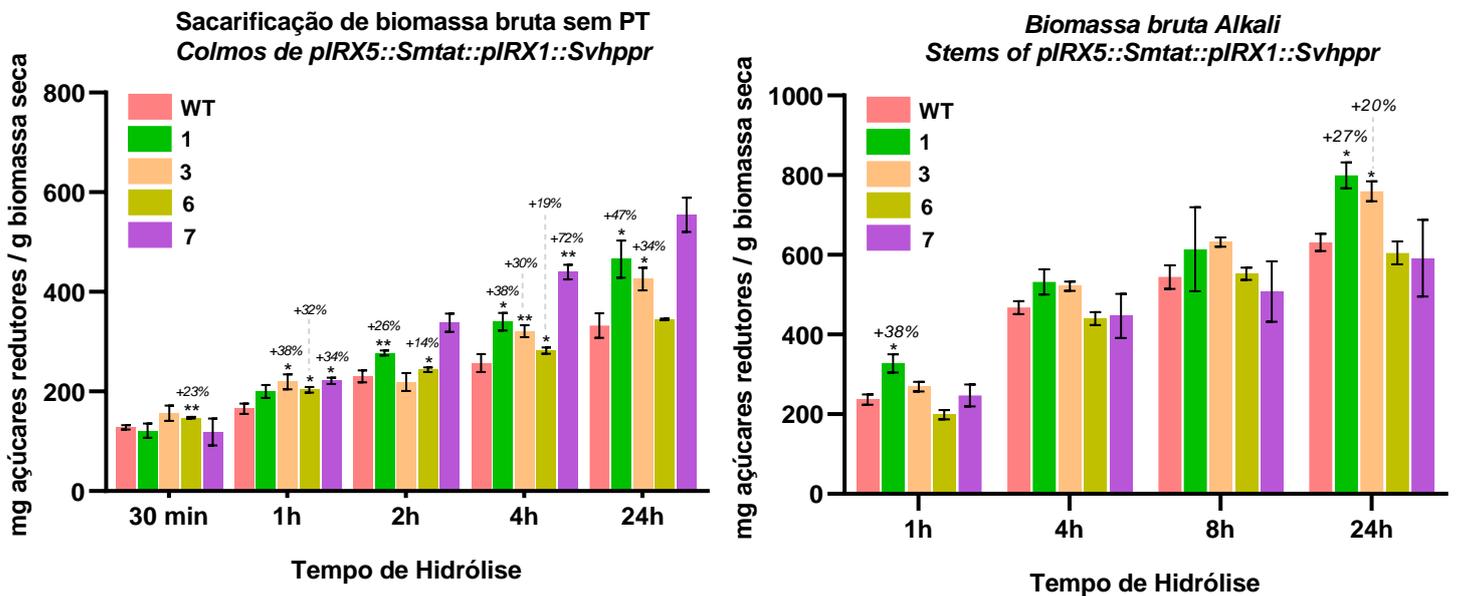
**Nota:** Cada linhagem é representada por pools, sendo pools A, B e C, totalizando um n = 3. Valores acima das barras indicando porcentagem de liberação de açúcar em comparação com WT para cada tempo de hidrólise. Barras de erro se referem ao erro padrão a partir da média (SEM). Asteriscos indicando significância estatística, sendo \*, P < 0.05; \*\*, P < 0.01; \*\*\*, P < 0.001; \*\*\*\*, P < 0.0001.

**Fonte:** do autor (2024).

Para a avaliação das linhagens de E432 foi realizado da mesma forma descrita anteriormente, avaliando-se a biomassa bruta e o CWR em condições sem pré-tratamento e após pré-tratamento alcalino com NaOH 50 mM. Foram avaliadas todas as linhagens obtidas, mas somente são demonstradas nas figuras as linhagens com resultados estatisticamente significativos. A Figura 14 demonstra a sacarificação da biomassa bruta sem pré-tratamento de linhagens de E432.

Observou-se que as linhagens 1, 3, 6 e 7 demonstraram maior liberação de açúcar em comparação com plantas selvagens, em diferentes tempos de hidrólise. Destaca-se, principalmente, a linhagem 7, com uma consistente maior liberação de açúcar nos tempos de leitura, com 72% mais açúcar liberado após 4h de reação; a linhagem 1, atingindo 47% mais liberação de açúcar após 24h; além das linhagens 3 e 6 com 38% e 32% mais liberação de açúcar após somente 1h, respectivamente.

**Figura 14** – Hidrólise enzimática de biomassa bruta de colmos de linhagens E432 em condição sem pré-tratamento e com pré-tratamento alcalino.



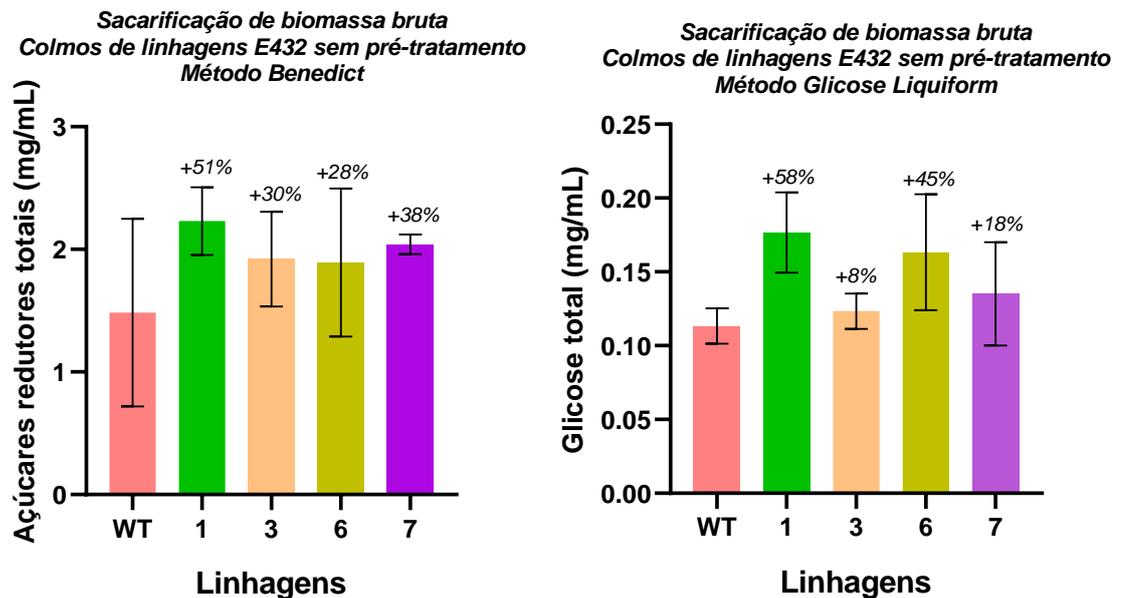
**Nota:** Cada linhagem é representada por pools, sendo pools A, B e C, totalizando um n = 3. A linhagem 7 possui apenas os pools A e B sendo para este n = 2. Valores acima das barras indicando porcentagem de liberação de açúcar em comparação com WT para cada tempo de hidrólise. Barras de erro se referem ao erro padrão a partir da média (SEM). Asteriscos indicando significância estatística, sendo \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ; \*\*\*\*,  $P < 0.0001$ .

**Fonte:** do autor (2024).

Como pode ser visto na avaliação das mesmas linhagens após pré-tratamento alcalino, foi possível verificar que as linhagens 1 e 3 demonstraram maior liberação de açúcar em comparação com plantas selvagens, mas apenas no tempo de hidrólise de 24h, sendo 27% para a linhagem 1 e 20% para a linhagem 3. As demais linhagens não demonstraram resultados estatisticamente significativos.

Como uma estratégia de confirmação dos resultados, utilizou-se a quantificação de açúcares por meio de outros métodos além do DNS, sendo a Glicose Liquiform para quantificação de glicose total, e reação de Benedict para quantificação de açúcares redutores. Por mais que os resultados não demonstrem significância estatística, observou-se que a tendência de aumento para todas as linhagens observadas anteriormente também é verdadeira utilizando-se estes métodos, como evidenciado na Figura 15.

**Figura 15** – Quantificação de açúcares pelos métodos Benedict e Glicose Liquiform.



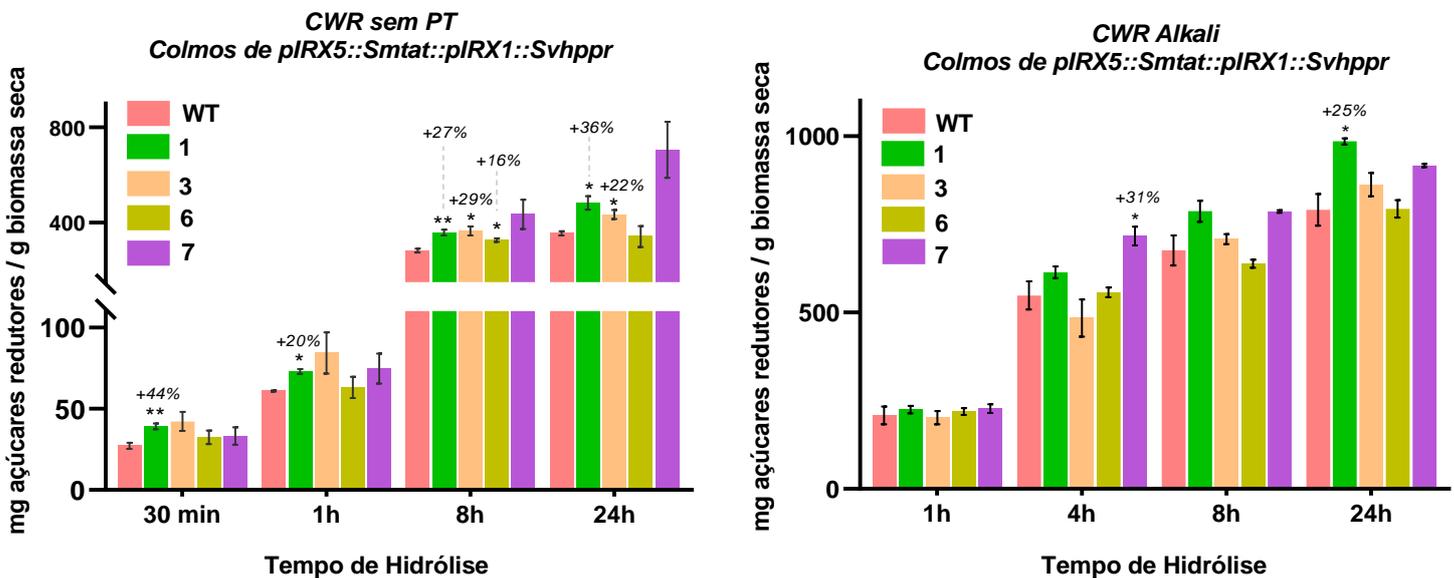
**Nota:** Cada linhagem é representada por pools, sendo pools A, B e C, totalizando um  $n = 3$ . A linhagem 7 possui apenas os pools A e B sendo para este  $n = 2$ . Valores acima das barras indicando porcentagem de liberação de açúcar em comparação com WT para cada tempo de hidrólise. Barras de erro se referem ao erro padrão a partir da média (SEM). Asteriscos indicando significância estatística, sendo \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ; \*\*\*\*,  $P < 0.0001$ .

**Fonte:** do autor (2024).

No caso da sacarificação utilizando CWR, verificou-se que na condição sem pré-tratamento as mesmas linhagens 1, 3, 6 e 7 mantiveram o perfil de maior liberação de açúcar, com resultados estatisticamente significativos para 1, 3 e 6, com destaque para a linhagem 1, que tem consistentes resultados estatísticos para todos os tempos de hidrólise. Já para a avaliação das mesmas linhagens utilizando CWR e com o pré-tratamento alcalino, verificou-se tendência de maior liberação de açúcar para as linhagens 1 e 7, onde a linhagem 1 demonstrou 25% mais liberação de açúcar após 24h, e a linhagem 7, demonstrou 31% maior liberação de açúcar no tempo de 4h, conforme pode ser visto na Figura 16.

Sendo assim, para as análises bioquímicas que serão apresentadas a seguir foram selecionadas as linhagens com perfil de sacarificação superiores às plantas selvagens, sendo as linhagens 1, 4, 6 e 8 da construção E430 e as linhagens 1, 3, 6 e 7 da construção E432.

**Figura 16** – Hidrólise enzimática de CWR de colmos de linhagens E432 sem pré-tratamento e com pré-tratamento alcalino.



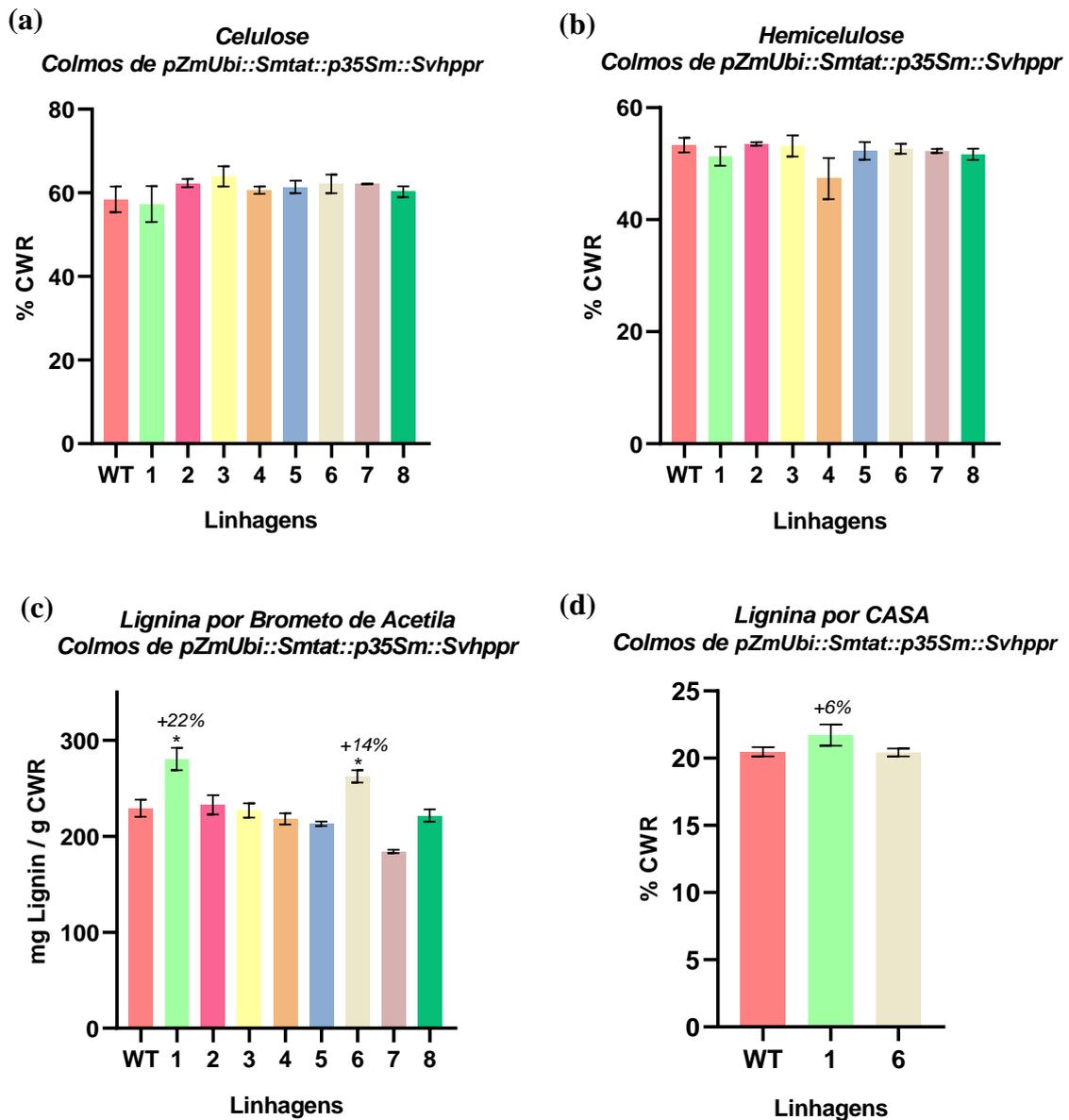
**Nota:** Cada linhagem é representada por pools, sendo pools A, B e C, totalizando um n = 3. A linhagem 7 possui apenas os pools A e B sendo para este n = 2. Valores acima das barras indicando porcentagem de liberação de açúcar em comparação com WT para cada tempo de hidrólise. Barras de erro se referem ao erro padrão a partir da média (SEM). Asteriscos indicando significância estatística, sendo \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ; \*\*\*\*,  $P < 0.0001$ .

**Fonte:** do autor (2024).

## 6.4 Composição da parede celular

Para os experimentos de conteúdo de lignina, celulose e hemicelulose utilizou-se o material da preparação da parede celular (CWR). Os resultados são apresentados em porcentagem de CWR, na Figura 17 para linhagens de E430, e na Figura 18 para as linhagens de E432. As Figuras incluem os experimentos de quantidade de celulose, hemicelulose, e lignina quantificada por dois métodos diferentes, brometo de acetila e CASA. Foram avaliadas todas as linhagens para cada construção para celulose e hemicelulose, e para lignina, o método CASA foi usado para confirmar os resultados das linhagens diferenciais pelo método de brometo de acetila.

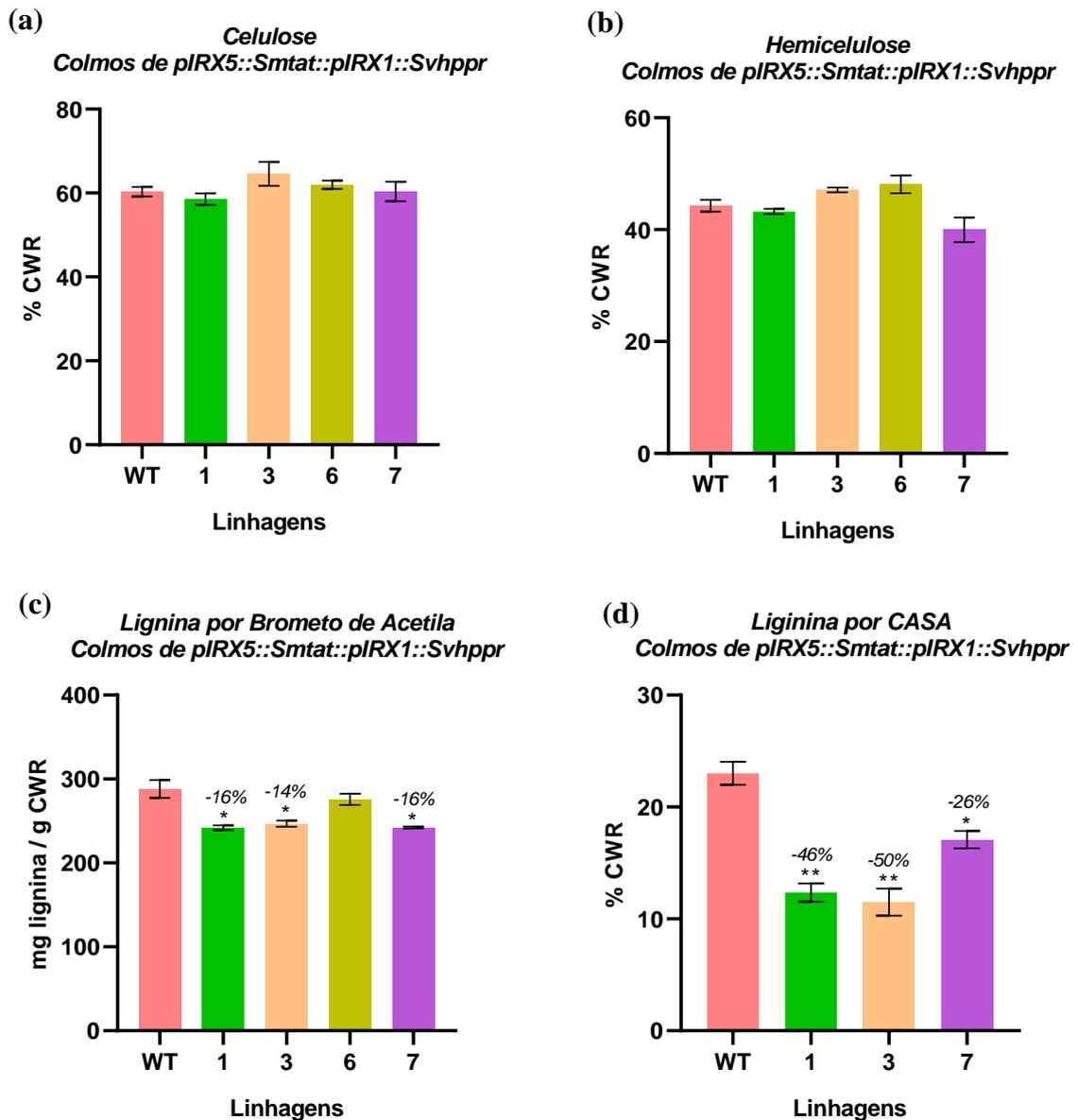
**Figura 17** – Caracterização da parede celular a partir do conteúdo de lignina, celulose e hemicelulose de linhagens de E430.



**Nota:** (a) Celulose pelo método do Reagente de Updegraff, sendo resultados expressos em porcentagem de CWR. (b) Hemicelulose após remoção por meio do uso de Ácido Trifluoroacético (TFA), e medida gravimétrica por diferença de peso inicial e final, expresso em porcentagem de CWR. (c) Lignina pelo método de Brometo de Acetila, sendo resultados expressos em porcentagem de CWR. (d) Lignina pelo método CASA – Cysteine-Assisted Sulphuric Acid. Linhagens representadas pelos pools A, B, C, totalizando um n = 3. Barras de erro se referem ao erro padrão a partir da média (SEM). Asteriscos indicando significância estatística, sendo \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ; \*\*\*\*,  $P < 0.0001$ .

**Fonte:** do autor (2024).

**Figura 18** - Caracterização da parede celular a partir do conteúdo de lignina, celulose e hemicelulose de linhagens de E432.



**Nota:** (a) Celulose pelo método do Reagente de Updegraff, sendo resultados expressos em porcentagem de CWR. (b) Hemicelulose após remoção por meio do uso de Ácido Trifluoroacético (TFA), e medida gravimétrica por diferença de peso inicial e final, expresso em porcentagem de CWR. (c) Lignina pelo método de Brometo de Acetila, sendo resultados expressos em porcentagem de CWR. (d) Lignina pelo método CASA – Cysteine-Assisted Sulphuric Acid. Linhagens representadas pelos pools A, B, C, totalizando um n = 3. Barras de erro se referem ao erro padrão a partir da média (SEM). Asteriscos indicando significância estatística, sendo \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ; \*\*\*\*,  $P < 0.0001$ .

**Fonte:** do autor (2024).

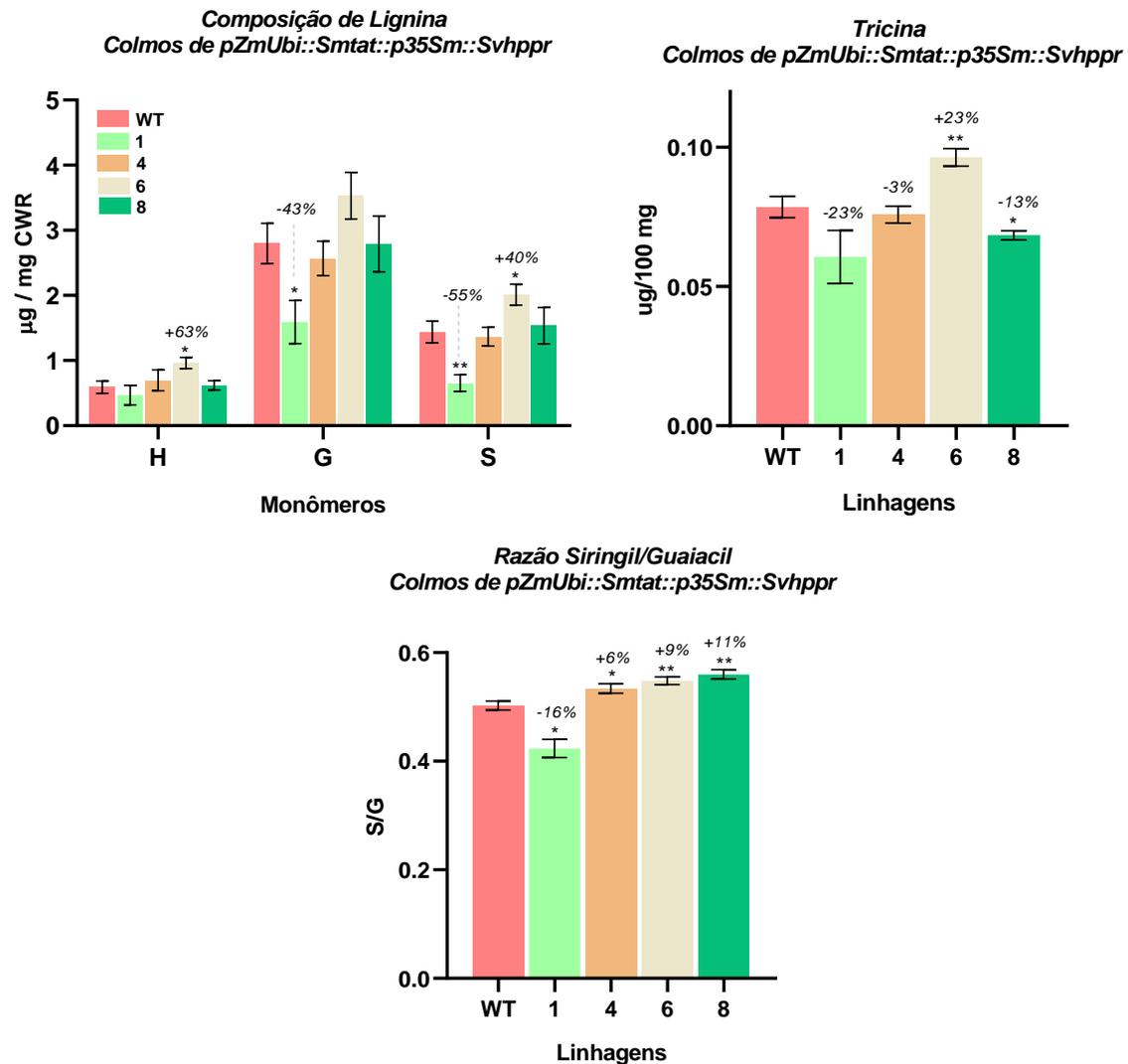
Conforme pode ser visto em **(a)** e **(b)** da Figura 17, para as linhagens de E430, o conteúdo de celulose e de hemicelulose das plantas transgênicas não foi alterado de maneira estatisticamente significativa em comparação com as plantas selvagens. Por outro lado, em **(c)** é possível evidenciar que o conteúdo de lignina das linhagens 1 e 6 se demonstrou alterado, a partir do método de brometo de acetila, em comparação com as plantas selvagens, sendo 22% maior para o evento 1 e 14% maior para o evento 6. Para confirmar este fato, realizou-se ainda a quantificação do conteúdo de lignina pelo método CASA, onde foi possível evidenciar uma tendência semelhante neste método, onde o evento 1 apresenta maior conteúdo de lignina que plantas selvagens.

Do mesmo modo, as linhagens da construção E432 evidenciaram não haver diferenças estatisticamente significativas para o conteúdo de celulose e também de hemicelulose. Mas, para o conteúdo de lignina, foi possível identificar um comportamento inverso ao que foi visto para E430: as linhagens de E432 demonstraram menor conteúdo de lignina para as linhagens 1, 3 e 7, e ainda mais relevante, de forma estatisticamente significativa por meio de dois métodos diferentes, brometo de acetila **(c)** e CASA **(d)**, como visto na Figura 18.

## 6.5 Composição de lignina por Tioacidólise

Uma vez definidas as linhagens com maior liberação de açúcares na sacarificação e/ou alteração no conteúdo de lignina, partiu-se para o experimento de tioacidólise para verificar se estas linhagens demonstravam alguma diferença na composição de lignina. Para a construção E430 foram avaliados os colmos das linhagens 1, 4, 6 e 8; e para a construção E432 foram avaliados os colmos das linhagens 1, 3, 6 e 7. Como pode ser visto na Figura 19, para a construção E430, foi possível evidenciar de forma estatisticamente significativa diferenças para os três monômeros avaliados, sendo uma diferença de 63% mais presença do monômero H e 40% do monômero S para a linhagem 6; e uma diferença de -43% do monômero G e -55% do monômero S para a linhagem 1. A razão S/G também foi avaliada, e evidenciou que a linhagem 1 tem razão diminuída em -16%; e as demais linhagens demonstraram que houve um aumento da razão S/G em comparação com plantas selvagens, sendo de 6% para a linhagem 4, 9% para a linhagem 6 e de 11% para a linhagem 8. Para a avaliação de tricina foi possível observar comportamentos diferentes entre as linhagens, sendo de uma redução para três linhagens, sendo uma estatisticamente significativa de -13% para a linhagem 8; e um aumento estatisticamente significativo de 23% para a linhagem 6.

**Figura 19** – Composição de lignina e razão S/G de linhagens E430.



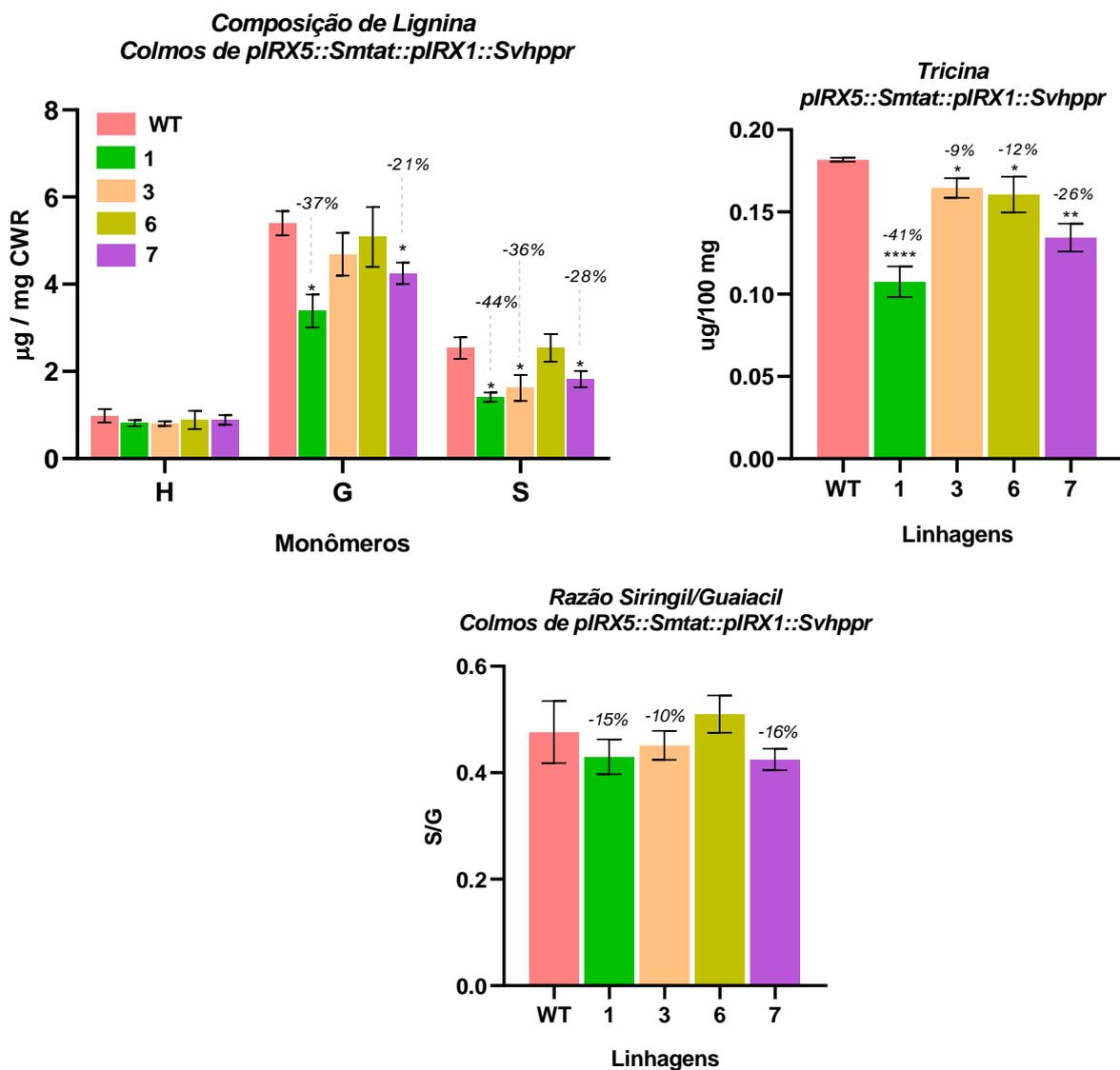
**Nota:** Linhagens representadas pelos pools A e B, em quadruplicata. Asteriscos indicando significância estatística, sendo \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ; \*\*\*\*,  $P < 0.0001$ . Os valores de erro apresentados se referem ao erro padrão a partir da média (SEM).

**Fonte:** do autor (2024).

Na avaliação das linhagens da construção E432, um comportamento de diminuição nos teores de cada um dos monômeros pôde ser observado, conforme evidencia a Figura 20. Neste caso, os resultados indicam menor participação dos monômeros G e S nas linhagens transgênicas, e valores bem próximos das plantas selvagens para o monômero H. A linhagem 1 demonstrou ter menores quantidades de G e S, sendo -37% e -44%, respectivamente; já a linhagem 3 demonstrou ter -36% de participação do monômero S na composição de lignina; e

a linhagem 7 demonstrou teores reduzidos de G, em -21%; e de S, sendo -28%, sendo todos estes resultados estatisticamente significativos. Para a avaliação da razão S/G, resultados confirmam uma tendência de redução, uma vez que os teores individuais de S e G diminuíram, era esperado que a razão S/G também indicasse ser menor que de plantas selvagens. Para a avaliação da tricina, observou-se que as linhagens apresentaram um comportamento linear, sendo uma menor presença de tricina de forma estatisticamente significativa para todas as linhagens, com destaque para linhagens 1 e 7, com -43% e -26%, respectivamente.

**Figura 20** – Composição de lignina e teores de S/G de linhagens E432.



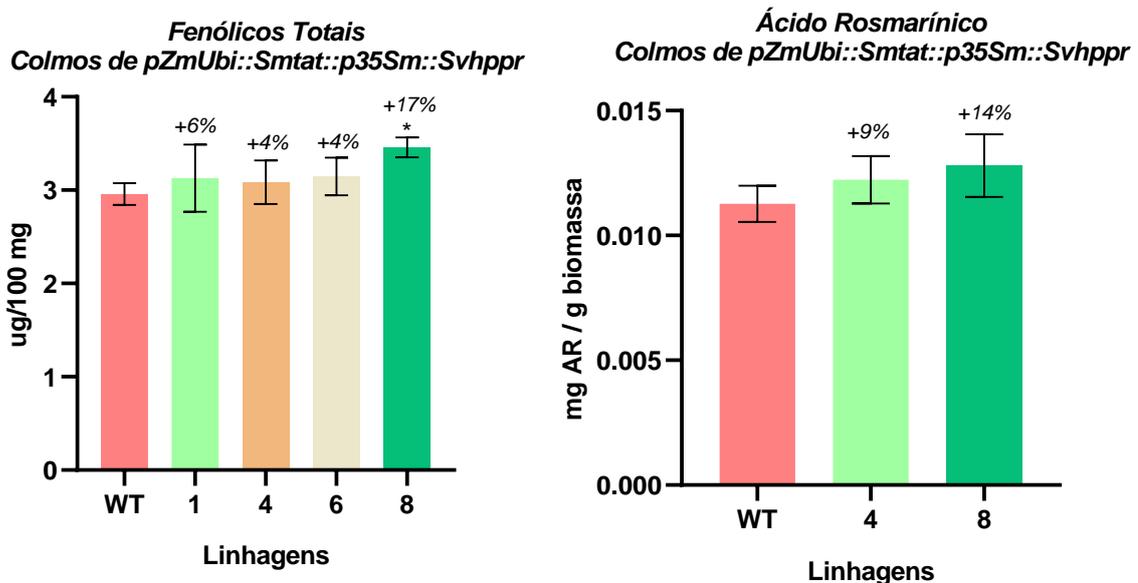
**Nota:** Linhagens representadas pelos pools A e B, em quadruplicata. Asteriscos indicando significância estatística, sendo \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ; \*\*\*\*,  $P < 0.0001$ . Os valores de erro apresentados se referem ao erro padrão a partir da média (SEM).

**Fonte:** do autor (2024).

## 6.6 Fenólicos Totais e Ácido Rosmarínico

Para a avaliação de fenólicos totais e a presença de ácido rosmarínico, foram analisadas as mesmas linhagens do experimento de tioacidólise, com o objetivo de verificar se houve diferenças estatisticamente significativas nos teores de fenólicos das plantas transgênicas, uma vez que o AR incorporado à parede celular poderia estar representado no teor total de fenólicos das amostras, portanto, esta análise poderia ser um indicativo da presença de AR ou da mudança no metabolismo fenólico das plantas, através da modificação genética. No caso dos colmos de linhagens da construção E430, observou-se que todas as amostras demonstraram uma tendência de maior presença de fenólicos totais que plantas selvagens, sendo estatisticamente significativo para a linhagem 8, representando 17% mais fenólicos que plantas WT. Para a presença de AR nos colmos, resultados indicaram uma tendência de maior presença nas linhagens transgênicas, porém, sem significância estatística, conforme a Figura 21. Também é importante mencionar que, como visto no Anexo III, Figuras 1 e 2, os padrões usados demonstraram comportamentos distintos, sendo observados em diferentes tempos de retenção, podendo ser algum efeito da coluna ou do método utilizado.

**Figura 21** – Fenólicos totais e AR de colmos da construção E430.

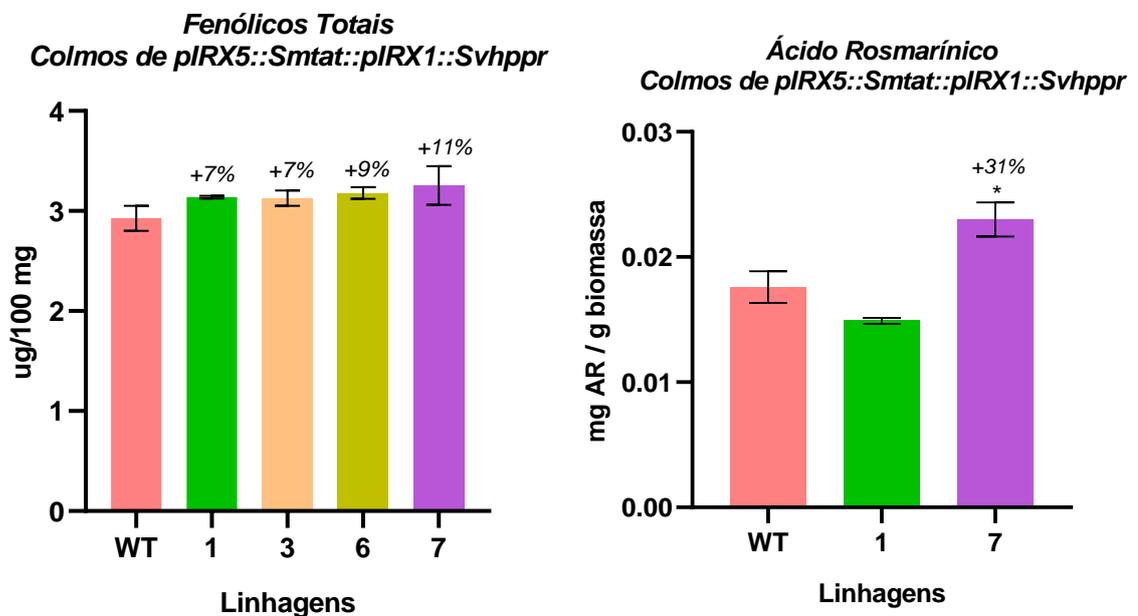


**Nota:** Linhagens representadas pelos pools A, B, C, totalizando um n = 3. Barras de erro se referem ao erro padrão a partir da média (SEM). Asteriscos indicando significância estatística, sendo \*, P < 0.05; \*\*, P < 0.01; \*\*\*, P < 0.001; \*\*\*\*, P < 0.0001.

**Fonte:** do autor (2025).

Para as linhagens de E432, os resultados dos experimentos de fenólicos totais das amostras das linhagens avaliadas demonstraram também uma tendência de mais fenólicos presentes nas linhagens transgênicas, mas sem resultados significativamente estatísticos. Além disso, na avaliação da presença de AR nos colmos destas linhagens, foi possível evidenciar que a linhagem 7 demonstrou resultados estatisticamente significativos para uma presença de AR 31% maior que em plantas selvagens, como pode ser visto na Figura 22. Os cromatogramas para os padrões e também para as amostras avaliadas são evidenciados no Anexo III – cromatogramas.

**Figura 22** – Fenólicos totais e AR de colmos da construção E432.



**Nota:** Linhagens representadas pelos pools A, B, C, totalizando um n = 3. Barras de erro se referem ao erro padrão a partir da média (SEM). Asteriscos indicando significância estatística, sendo \*, P < 0.05; \*\*, P < 0.01; \*\*\*, P < 0.001; \*\*\*\*, P < 0.0001.

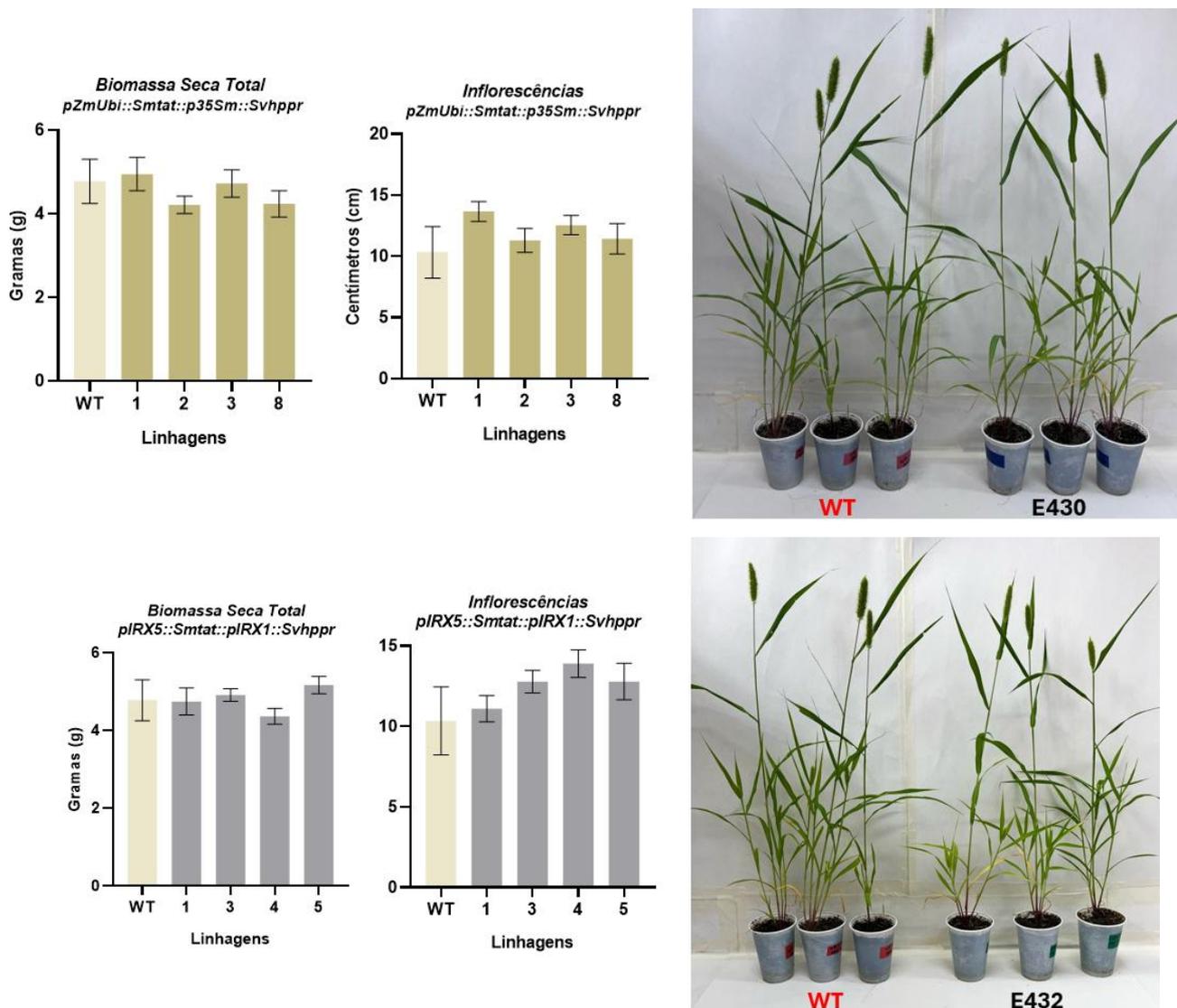
**Fonte:** do autor (2025).

## 6.7 Biomassa e Inflorescências

Na avaliação de biomassa seca total e inflorescências, não foram observadas diferenças estatisticamente significativas para nenhum dos parâmetros, em nenhuma das linhagens, para as duas construções. Este fato é importante pois indica que a modificação genética não afetou as linhagens transgênicas nos parâmetros de biomassa e de inflorescências, com diferentes

linhagens demonstrando o mesmo comportamento de crescimento que plantas selvagens, sem nenhuma distinção fenotípica. Além disso, a similaridade fenotípica com as plantas selvagens, como visto na Figura 23, também indica que as linhagens transgênicas se adaptaram às suas modificações na parede celular, como teores alterados de lignina, a sua composição e também de compostos fenólicos que podem estar associados ao metabolismo fenólico e a biossíntese de lignina.

**Figura 23** – Biomassa e Inflorescências de linhagens E430 e E432.



**Nota:** Nas fotos, WT a esquerda, e linhagens da construção E430 e E432 a direita. Linhagens representadas por plantas individuais, totalizando um  $n = 8$  para cada linhagem. Biomassa seca incluindo inflorescências e sementes. Inflorescências em formação também foram consideradas. Barras de erro se referem ao erro padrão a partir da média (SEM). Asteriscos indicando significância estatística, sendo \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ; \*\*\*\*,  $P < 0.0001$ .

**Fonte:** do autor (2025).

## 6.8 Metabolômica

Para os resultados obtidos por metabolômica, priorizou-se o modo negativo onde os fenólicos são detectados com maior sensibilidade e seletividade (Gouveia-Figueira, Castilho, 2015), e foram consideradas as linhagens 4 e 8 representando a construção E430 e as linhagens 1 e 7 representando a construção E432. Como pode ser visto a seguir na Tabela 2, alguns compostos fenólicos envolvidos tanto na biossíntese de AR quanto de lignina, demonstraram perfis diferentes do perfil de plantas selvagens. Na Tabela 2, podemos observar os resultados e a sua respectiva diferença percentil em comparação com plantas selvagens, para colmos e para folhas. Apesar de alguns compostos não terem sido detectados, como AR e outros da via de lignina, podemos ver fenólicos importantes para estas vias como os ácidos caféico, ferúlico, p-cumárico, tricina, tirosina, e outros compostos relacionados sendo alterados de forma negativa nas plantas transgênicas. Ou seja, alinhando-se estes resultados com os resultados anteriores, principalmente aqueles que indicam por exemplo menor conteúdo de tricina por tioacidólise, e modificação da composição e quantidade de lignina, é correto afirmar que o metabolismo de fenólicos das plantas transgênicas foi alterado, e a alteração se deu principalmente em compostos das vias de lignina, onde os indícios apontam para uma regulação negativa de diferentes moléculas envolvidas nestes processos.

É possível ver ainda que estas modificações podem causar diferentes comportamentos, como por exemplo o glicosídeo do ácido p-cumárico, que não foi detectado em colmos para nenhuma construção, mas demonstrou aumento importante para folhas de E430, e redução estatisticamente significativa para folhas de E432, sendo assim é possível verificar que colmos e folhas podem apresentar comportamentos distintos. A detecção dos compostos foi mais evidenciada em folhas do que em colmos para algumas moléculas, e isto pode ser justificado pelo método de extração, que pode ter sido mais eficiente para tecidos de folhas. Além disso, é importante mencionar que o método de extração - que foi definido pelo centro de pesquisa que realizou a análise - também pode ser responsável pela não detecção de AR, uma vez que é um método diferente do que foi usado para a extração para identificação de AR por HPLC.

**Tabela 2** – Anotação de possíveis compostos diferenciais de E430 e E432.

Composto	PARÂMETROS		COLMOS		FOLHAS		REFERÊNCIA
	RT	<i>m/z</i>	E430	E432	E430	E432	
Ácido p-cumárico	6.69	163.0399	0	0	17%	*-87%	Bondia-Pons <i>et al.</i> , (2014)
Coniferil aldeído	7.61	177.0557	-94%	*-85%	-65%	*-89%	Prothmann, Jens <i>et al.</i> , (2020)
Ácido cafeico	10.64	179.0564	-59%	71%	-58%	-54%	Zhao <i>et al.</i> , (2019)
Ácido ferúlico	8.23	193.0506	0	0	-12%	*-87%	Zhao <i>et al.</i> , (2019)
Tricina	14.33	329.0665	-77%	53%	-44%	-77%	Composto padrão
Ácido clorogênico	21.64	353.2085	*22%	*-25%	-36%	-71%	Ye <i>et al.</i> , (2012)
L-tirosina	8.58	180.9729	19%	15%	-16%	*-84%	Yamamoto <i>et al.</i> , (2023)
1-O-trans-feruloil- $\beta$ -glucopiranosídeo	8.23	355.1037	-64%	-77%	49%	*-81%	Alarcón <i>et al.</i> , (2021)
Glicosídeo ácido p-cumárico	7.77	325.0931	0	0	176%	** -94%	Bondia-Pons <i>et al.</i> , (2014)
Ácido 1-ferulol-quinico	7.90	367.1038	-100%	-100%	-24%	-86%	Grayer <i>et al.</i> , (2000)
Feruloil tiramina	21.64	312.1720	-15%	-31%	-35%	-77%	Jiang <i>et al.</i> , (2021)
Ácido p-hidroxi-benzoico	21.67	144.9240	-21%	*-31%	-31%	-34%	Yan, Kaiser, (2018)

**Nota:** Anotação de compostos baseados somente nos valores de *m/z* observados na literatura, sem levar em conta os espectros de fragmentação. Linhagens 4 e 8 representando construção E430 e linhagens 1 e 7 representando construção E432, considerando-se a média das diferenças das linhagens em relação às plantas selvagens, a partir dos valores de abundância. Resultados representados pelos pools A, B e C de cada linhagem, em triplicata. Resultados representados de forma percentil representam diferença das linhagens em comparação com plantas selvagens para cada composto. Barras de erro se referem ao erro padrão a partir da média (SEM). Asteriscos indicando significância estatística, sendo \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ; \*\*\*\*,  $P < 0.0001$ .

**Fonte:** do autor (2025).

## 7. Discussão

Sumarizando-se os resultados, e considerando para esta discussão somente as linhagens com resultados mais homogêneos, sendo as linhagens 4 e 8 para E430 e as linhagens 1 e 7 para E432, temos o seguinte cenário: no caso das linhagens de E430, utilizando promotores constitutivos, foi evidenciada uma maior liberação de açúcar somente para a sacarificação sem pré-tratamento utilizando CWR; além disso, na avaliação da parede celular os conteúdos de celulose e hemicelulose não demonstraram modificações; assim como para lignina, que demonstrou um maior conteúdo para uma das linhagens, somente por um dos métodos utilizados, mostrando-se bem similar à de plantas selvagens. A lignina destas linhagens também se demonstrou alterada em sua composição de forma estatisticamente significativa para tricina. Por fim, a quantidade de fenólicos totais e de AR demonstraram tendência de aumento em

comparação com as plantas selvagens. No caso das linhagens de E432, utilizando-se promotores ativos somente na formação da parede celular secundária, observamos maior liberação de açúcar para experimentos sem pré-tratamento e também após pré-tratamento alcalino, tanto para biomassa bruta como para CWR; sem modificações no conteúdo de celulose e hemicelulose; e para a lignina, ao contrário da outra estratégia, observou-se teores menores no conteúdo total em comparação com plantas selvagens, assim como para os monômeros S e G, além da tricina, todos reduzidos. O teor de fenólicos totais e de AR também demonstrou uma tendência de maior presença nas plantas transgênicas em comparação com selvagens. Para a metabolômica, nas linhagens transgênicas foi possível evidenciar quantidades reduzidas em comparação com as selvagens, em diferentes compostos que estão envolvidos no metabolismo fenólico das plantas, para ambas as construções.

### **7.1 A eficiência de transformação da construção com promotores constitutivos**

A partir da superexpressão dos genes *SmTAT* e/ou *SvHPPR* envolvidos na biossíntese de ácido rosmarínico, e da possível modificação da composição da lignina, este trabalho buscou verificar se a recalcitrância da biomassa lignocelulósica seria diminuída, e ao mesmo tempo, se a hidrólise enzimática teria a eficiência aumentada, sem afetar o desenvolvimento das plantas. De acordo com os resultados evidenciados anteriormente, alguns pontos positivos muito interessantes que suportam as hipóteses aqui testada devem ser considerados.

Primeiramente, a partir dos resultados de genotipagem, é possível concluir que o protocolo de transformação genética utilizado foi eficiente do ponto de vista de obtenção de linhagens transgênicas, mesmo que a eficiência tenha sido menor que a relatada na literatura. O protocolo de transformação proposto por Martins *et al.*, (2015) é bastante simples, embora seja necessário a realização de várias etapas de trocas de meio de cultura. Os autores relatam uma eficiência de transformação na ordem de 29%, utilizando *phytotron* para crescimento das plantas. Neste trabalho, a quantidade de transformações para obter um número suficiente de plantas transgênicas para as análises fenotípicas foi bastante alta, chegando a mais de 40 transformações para obtenção de 18 eventos confirmados. Além disso, foi notória uma maior dificuldade na transformação da construção E430, utilizando promotores constitutivos, em relação à outra construção, de promotores de parede celular secundária. Notava-se que os explantes chegavam até a etapa de regeneração de raízes com o meio MRS, mas somente alguns

poucos conseguiram avançar até a etapa de crescimento de fase aérea com o uso do meio MD. Uma hipótese para explicar este fenômeno é a de que o acúmulo de AR, ou outros metabólitos, estaria sendo tóxico a ponto de não permitir o desenvolvimento dos explantes, uma vez que o promotor utilizado atua de maneira constitutiva, ou seja, ativo em todos os tecidos das plantas. Para a confirmação desta hipótese, poderiam ser realizados experimentos utilizando-se os tecidos ainda da fase de regeneração, para verificar o conteúdo de fenólicos totais e de AR.

Corroborando esta hipótese, os autores Bevilaqua *et al.*, (2019) mostraram que em plantas de milho a aplicação exógena de AR tem efeitos deletérios nos primeiros estágios de desenvolvimento, causando consequências negativas, o que pode indicar toxicidade do composto. Os autores também citam que a aplicação de AR inibe duas enzimas relacionadas com a biossíntese de lignina, PAL e TAL, extremamente importantes para o desenvolvimento das plantas, o que também pode estar relacionado. A escolha dos promotores IRX, ativos somente durante a formação da parede celular, teve justamente o objetivo de induzir o aumento da expressão dos genes somente apenas após o desenvolvimento inicial das plântulas. Para os 10 eventos transgênicos obtidos para a construção E432, não se observou a mesma dificuldade que para E430, porém, também foi evidenciada uma baixa eficiência de transformação, devido aos problemas de manuseio, condições de crescimento e contaminação inerentes aos experimentos de transformação genética e cultura de tecidos.

Na etapa de avaliação da expressão gênica, evidenciou-se que as linhagens obtidas demonstraram superexpressão variada de *SvHPPR*, e algumas linhagens não demonstraram ou demonstraram uma expressão baixa (se comparado à *SvHPPR*) do transgene *SmTAT*, e isto foi verdadeiro para as duas construções. Este fato pode estar atrelado a fatores relacionados à integração dos transgenes no genoma das plantas. Uma vez que a integração dos transgenes é aleatória, diferentes linhagens podem apresentar comportamentos diferentes no nível de expressão, onde a expressão dos genes pode ser influenciada por diversos fatores, como por exemplo a integração em orientação invertida, o que faz com que os transgenes sejam transcritos de forma convergente; ou ainda a integração de uma cópia única de T-DNA, que também pode resultar baixas ou nenhuma expressão dos genes; ou também, pode ser explicado por mecanismos de controle que estão sendo ativados nas plantas a partir da inserção do T-DNA, implicando em supressão e silenciamento dos transgenes (DE BUCK; DE PAEPE; DEPICKER, 2013).

São necessárias maiores investigações para explicar o motivo da baixa ou nula expressão de *SmTAT* nas linhagens avaliadas, porém, como visto na seção 6.1 dos Resultados,

a espécie *S. viridis* não tem homólogos do gene *SmTAT*, sendo assim, a sequência de *S. miltiorrhiza* utilizada poderia estar impedindo a correta funcionalidade dos sistemas de transcrição e tradução. Por outro lado, as evidências demonstram que a expressão de *SvHPPR* foi aumentada de forma significativa em quase todos os eventos avaliados. Isto confirma a capacidade do promotor 35S utilizado, neste caso, otimizado para o uso em monocotiledôneas. O promotor 35S também é conhecido na literatura pela sua aptidão de garantir elevada transcrição, não sendo muito influenciado por condições ambientais ou tipo de tecido. No entanto, também é sabido que a expressão de transgenes dirigidos por 35S pode ser bastante variável em um grupo de linhagens transgênicas com a mesma construção, de acordo com De Buck; De Paepe; Depicker (2013), o que pode explicar a diferença dos níveis de expressão entre as linhagens de uma mesma construção.

No caso dos eventos da construção E432 a escolha dos promotores utilizados teve como objetivo a ativação dos genes somente durante a formação da parede celular secundária dos vasos do xilema, que se forma apenas quando a planta atinge a maturidade. Em *Arabidopsis thaliana*, os genes *IRX1* e *IRX5* foram reportados na literatura como sendo essenciais para a formação da parede celular secundária. Estes genes são subunidades da celulose sintase (CesA), sendo *IRX1* – CesA8 e *IRX5* – CesA4, dois dos três genes CesA responsáveis pela síntese de celulose nas paredes celulares secundárias dos vasos do xilema em desenvolvimento, que ocorre depois da cessão da expansão celular (HERNÁNDEZ-BLANCO *et al.*, 2007); (GARDINER; TAYLOR; TURNER, 2003).

Sendo assim, utilizou-se as regiões à montante dos genes, correspondentes às regiões dos promotores de cada um, ativos somente durante a formação da parede celular secundária dos vasos do xilema, mais presentes nos colmos das plantas, mas também presente em folhas. Foi possível evidenciar que os promotores têm um ótimo funcionamento em *S. viridis*, como visto na Figura 11, para a expressão gênica em colmos de plantas na fase *pré-booting*, e também pelo experimento de fases, evidenciado no Anexo V, onde foi possível verificar que os promotores já estavam ativos em plantas com 15 e 30 dias de crescimento.

## **7.2 A maquinaria de AR e a presença de TAT em *S. viridis***

No estudo da rota de biossíntese de AR em *S. viridis* a partir da ferramenta BLAST, foi possível observar as enzimas *PAL*, *C4H*, *4CL*, *TAT*, *RAS*, *HPPR*, *CYP98A* com médios a altos valores de identidade e também baixos valores de *e-value*, indicando que a probabilidade de a

similaridade observada ser devido ao acaso é baixa, para maioria dos casos. No caso de *TAT*, observou-se não haver genes homólogos, e o *e-value* mais satisfatório dos hits encontrados ainda é considerado alto em comparação com os demais, justificando a escolha pelo gene de *S. miltiorrhiza* para construção dos vetores de expressão. Portanto, sugerimos que *S. viridis* poderia ser capaz de produzir AR, uma vez que a espécie possui grande parte das enzimas chave do seu processo de biossíntese. Além disso, os resultados da análise de HPLC para identificação de AR indicam que esta maquinaria estaria funcionando, uma vez que a molécula de AR foi possivelmente identificada tanto em plantas selvagens, quanto em plantas transgênicas. Porém, a confirmação desta hipótese só poderia vir de uma análise mais detalhada – como por espectrometria de massas – para confirmar se a molécula identificada é mesmo de AR. Além disso, seria necessário verificar se a produção de AR estaria acontecendo através da ação do gene *TAT* endógeno, de *S. viridis*, ou do gene *SmTAT*, utilizado no vetor de expressão. A hipótese mais provável é que a maquinaria de biossíntese de AR esteja funcionando a partir do transgene *SmTAT* para as linhagens 4 e 8 de E430, onde observou-se a presença de AR solúvel a partir da análise por HPLC, e também se observou superexpressão do gene na análise de expressão gênica de *SmTAT*. Para as linhagens 1 e 7 de E432, não se verificou aumento da expressão de *SmTAT*, mas verificou-se a presença de AR solúvel, o que poderia indicar que, neste caso, a maquinaria de biossíntese de AR poderia estar funcionando a partir do gene *TAT* endógeno de *S. viridis*.

### 7.3 Limitações do método de sacarificação

Com visto na seção 6.3 dos Resultados, a sacarificação evidenciou linhagens com perfis de liberação de açúcar diferentes de plantas selvagens, com algumas linhagens demonstrando maior liberação de açúcar. Porém, é necessário considerar alguns aspectos relacionados aos experimentos de sacarificação. O primeiro deles é o fato de que utilizando o material de CWR as diferenças na hidrólise aparecem de forma estatisticamente significativa somente após 24h de reação, muito depois do que acontece quando se utiliza o material da biomassa bruta, como visto nas Figuras 12 e 13 para E430. É importante mencionar que no caso da biomassa bruta é possível encontrar outros açúcares não-estruturais, como por exemplo amido, xilose, arabinose e glicose solúveis que estão presentes na biomassa, mas que podem ter sido retirados na lavagem da parede celular, no protocolo de preparação do CWR. Na literatura, milhetos como *S. viridis* são descritos como ricos em amido (Ermawar *et al.*, 2015), e isto poderia ser a

explicação para a rápida liberação de açúcar da biomassa bruta, influenciada pela presença do amido e outros açúcares não-estruturais.

Outro aspecto importante sobre a sacarificação é relacionado ao fato de que os experimentos utilizando pré-tratamento alcalino não terem demonstrado perfil de maior liberação de açúcar para as linhagens da construção E430, onde não foram observadas diferenças para o material de CWR, utilizando-se NaOH 50 mM. Este fato poderia estar relacionado ao fato de as linhagens desta construção apresentarem maior quantidade de lignina, como mostrado na Figura 17. A quantidade de lignina afeta diretamente a eficiência do pré-tratamento alcalino e da hidrólise, com maiores teores de lignina dificultando o acesso aos açúcares. Outros fatores também são interferentes importantes como as propriedades e a localização da lignina na biomassa, as condições da superfície da lignina e a sua área de superfície; além disso a ligação não produtiva da enzima na lignina também pode contribuir para uma baixa eficiência, uma vez que a enzima fica inativa, sendo este aspecto diretamente relacionado com a superfície e a localização da lignina (KIM, LEE, KIM, 2016); (YANG, WYMAN, 2006).

Portanto, a mudança da estrutura da lignina, como visto nos resultados de tioacidólise, aliada ao aumento da quantidade da lignina, visto nos experimentos de brometo de acetila e CASA, na parede celular, pode estar contribuindo para uma menor eficiência dos pré-tratamentos para linhagens de E430. Já para linhagens de E432, a biomassa bruta e o CWR tratados com NaOH demonstraram maior liberação de açúcar após 4h e 24h, para linhagens com menores quantidade de lignina, corroborando a hipótese citada anteriormente.

#### **7.4 Mudanças na composição da lignina e a possível influência do Ácido Rosmarínico na parede celular das plantas transgênicas**

Uma vez identificadas as linhagens com maior liberação de açúcar, foi possível direcionar o foco das demais avaliações para estas linhagens, concentrando-se em como essas plantas se diferenciavam de plantas selvagens. Considerando-se conteúdo de lignina como um parâmetro essencial a ser avaliado, observou-se que para as linhagens E430, a partir do uso de promotores constitutivos, obteve-se quantidades de lignina muito semelhantes à de plantas selvagens, na avaliação da lignina solúvel pelos métodos brometo de acetila e CASA. Importante destacar que ambos são métodos indiretos, mas diferentes para avaliação de lignina,

sendo o brometo de acetila um método para avaliação de lignina solubilizada, onde produtos furfural derivados de xilanos podem interferir na absorvância a 280 nm; enquanto o método CASA utiliza-se da completa dissolução da biomassa com cisteína e ácido sulfúrico para quantificação da lignina observada também em 280 nm, podendo sofrer da mesma interferência, e, sendo métodos considerados inferiores ao método Klason (Lu *et al.*, 2021), cujo protocolo ainda não estava estabelecido no laboratório.

Na avaliação da composição de lignina obtida a partir da tioacidólise, observamos que a composição de lignina das linhagens 4 e 8, sob o controle de promotores constitutivos, é bastante similar às plantas selvagens, tanto para o monômero G quanto para o monômero S, mas revelou diferenças estatisticamente significativas para tricina, no caso da linhagem 8, e para a razão S/G de todas as linhagens observadas em comparação com plantas selvagens. Em um estudo de Kishimoto *et al.*, (2005) a lignina com alta razão S/G foi relatada como mais linear, e com maior presença do monômero S. Análises *in silico* de Vermaas *et al.*, (2019) mostraram que em casos de lignina com alta razão S/G as energias livres produzidas pela exclusão aquosa dos grupos metoxi em monômeros S podem promover uma associação do polímero de lignina com a matriz de celulose de forma mais próxima, contribuindo para um mecanismo de exclusão de água que protege os carboidratos da parede celular da exposição e degradação por ácidos de origem patogênica, ROS e enzimas degradadoras da parede celular, o que poderia dificultar a hidrólise enzimática (PERACCHI *et al.*, 2024).

A razão S/G e a abundância relativa dos monômeros principais de lignina sempre foram relacionados com a hidrólise enzimática na literatura, porém, diferentes relações foram descritas ao longo dos anos. Os autores Studer *et al.*, (2011) relatam que um alto teor de S/G poderia ser benéfico para a liberação de glicose e xilose em amostras após pré-tratamentos; o que não é o caso do nosso estudo para as linhagens E430, com alto S/G, mas sacarificação ineficiente após pré-tratamentos. No entanto, outros estudos sugerem que alta relação de S/G está relacionada negativamente com a eficiência da hidrólise enzimática após pré-tratamentos (Tan *et al.*, 2015), o que está de acordo com os resultados encontrados neste estudo para linhagens E430, com razão S/G maior que de plantas selvagens. Isto pode ser visto no Anexo V, para a sacarificação após pré-tratamento alcalino com NaOH de CWR indicando perfis de hidrólise enzimática bastante semelhantes às de plantas selvagens.

Observando-se os resultados para fenólicos totais e presença de AR destas linhagens de E430 com promotores constitutivos, podemos perceber que há uma maior presença de compostos fenólicos nas plantas transgênicas, assim como a presença de AR, que pode estar

atuando de alguma forma no metabolismo fenólico e na biossíntese de lignina. Além disso, os resultados da metabolômica de colmos e de folhas para estas linhagens de promotores constitutivos reforçam que há alterações na regulação do metabolismo fenólico destas plantas em comparação com plantas selvagens. Segundo Xiao *et al.*, (2011) a expressão dupla de *TAT-HPPR* produz o maior conteúdo de AR, além de metabólitos associados a lignificação, como ácidos fenólicos. Os autores destacam que a expressão conjunta destes genes causa uma mudança no fluxo metabólico, modificando a via de produção de AR, diminuindo os níveis de ácido homogentístico e aumentando o acúmulo de ácidos fenólicos. Isto está de acordo com os resultados observados neste estudo, onde foi possível verificar um teor de fenólicos maior que para plantas selvagens, e levando-se em conta os resultados de expressão gênica, podemos ver que as linhagens 4 e 8 se destacam para a expressão conjunta dos dois genes. Notavelmente, os resultados de compostos fenólicos para as linhagens sob controle de promotores constitutivos (E430) são contrastantes em relação àqueles observados em plantas expressando os genes alvo sob controle dos promotores de parede celular secundária (E432). Sendo assim, sugere-se que plantas transgênicas superexpressando os genes alvo com promotores constitutivos, como no caso de E430, conseguem se adaptar para um correto balanceamento da produção e composição de lignina na parede celular, uma vez que os promotores estão ativos desde o início da deposição e polimerização da lignina, resultado em modificações menos acentuadas, como visto nos resultados apresentados.

Para corroborar esta hipótese, levamos em conta os resultados obtidos com as linhagens usando promotores de parede celular secundária, E432, onde houve uma diminuição no conteúdo de lignina, alteração mais acentuada da composição da lignina e regulação negativa de vários compostos fenólicos que fazem parte da rota de biossíntese de lignina, como visto na Tabela 2 dos resultados de metabolômica.

Para os resultados vistos na avaliação das linhagens utilizando-se promotores de parede celular secundária, de E432, podemos observar diferenças no conteúdo de lignina – reduzidos, e principalmente na sua composição. Para a composição de lignina, verificou-se que os monômeros G e S, além da tricina, demonstraram menor participação na composição da lignina em plantas transgênicas, resultando em uma menor razão S/G e em menor quantidade de lignina na parede celular. A tricina é formada a partir da via de fenilalanina, através de p-cumaroil-CoA, que é convertido através de p-cumarato, que é um braço comum para os três monolignóis, pela ação da enzima 4-cumarato:CoA ligase (4CL) (COOMEY, *et al.*, 2020). Este intermediário atua como substrato inicial para a via, e pode ser direcionado para a biossíntese de tricina. A

unidade p-cumaroil-CoA também participa da biossíntese da unidade G, fornecendo feruloil-CoA, que é convertido em coniferilaldeído, e coniferil álcool até a formação da unidade G (EUDES *et al.*, 2017).

Sendo as vias PAL e TAL diretamente relacionadas com a síntese de p-cumarato (Maeda, 2016), e por fim, p-cumaroil-CoA, quaisquer perturbações nessas vias poderiam alterar as quantidades de S, G e tricina. Na literatura, inclusive, sugere-se que a atividade de TAL em gramíneas poderia influenciar a atividade de C4H, contornando-a para a produção de cumarato (COOMEY, *et al.*, 2020). Como toda esta via precursora de p-cumaroil-CoA é proveniente da via de tirosina, é sugerido que perturbações geradas nesta via estejam direcionando o ácido p-cumárico formado para outras vias, resultando em menores quantidades de tricina, e dos monômeros G e S, uma vez que o substrato inicial da via, a tirosina, está sendo compartilhado também para a formação de AR na sua via de biossíntese. Para corroborar este fato, podemos ver resultados de metabolômica indicando que muitos compostos fenólicos estão reduzidos nas plantas transgênicas, o que poderia ser mais um indício do compartilhamento dos precursores.

Desta forma, o que se sugere é que o compartilhamento de precursores das vias de lignina e AR seja responsável pelo redirecionamento dos substratos para a via de AR ao invés da formação de S, G e tricina. Isto pode ser verdadeiro caso a expressão dos genes esteja ocorrendo somente após a deposição da parede celular secundária, a partir dos promotores utilizados. Para corroborar este fato, experimentos de expressão gênica de linhagens de E432 (Anexo V – Figura 1) indicam que os promotores estão ativos a partir de 15 dias de crescimento, e permanecem com um grau de expressão ainda maior para plantas com 30 dias de crescimento, quando a parede celular já foi depositada. Uma vez que as células já se encontram maduras, a deposição celular encerra a expansão celular, e as células percebem que já estão prontas para o armazenamento de açúcares, sendo então enriquecida com transportadores de açúcar, fotossíntese, e biossíntese de flavonoides, como no caso da tricina, conforme descreve Martin *et al.*, (2016). Assim, sugere-se que o metabolismo fenólico foi muito mais afetado em plantas transgênicas sob controle dos promotores de parede celular secundária (*IRX1* e *IRX5*), especulando-se que estas plantas podem não ter tido tempo suficiente para regular tal metabolismo, em contraste com plantas sob controle de promotores constitutivos.

De maneira geral, os resultados observados em nossos estudos sugerem que a estratégia de engenharia genética utilizada é promissora para alteração da composição da lignina em gramíneas. Uma estratégia similar realizada em uma dicotiledônea foi descrita na literatura utilizando a superexpressão de genes envolvidos na biossíntese de Escopoletina em plantas de

*A. thaliana*, com o mesmo objetivo de reduzir a recalcitrância da biomassa através do acoplamento de moléculas. No estudo, teores reduzidos do monômero G são relatados, além de melhora da sacarificação dos colmos das plantas. Os autores sugerem que a Escopoletina estaria se acoplando aos monolignóis, principalmente na posição O-4, por diferentes métodos de acoplamento, e atuando como uma unidade inicial do polímero da lignina (Hoengenaert, *et al.*, 2022). Da mesma forma, o AR também poderia estar atuando como uma unidade inicial da lignina, uma vez que os teores de S, G e tricina estão diminuídos devido ao possível compartilhamento dos precursores da rota de biossíntese. Porém, é importante destacar que este trabalho utilizou a dosagem de AR na forma solúvel a partir da biomassa bruta, e o mesmo composto não foi identificado na metabolômica. Dessa forma, não é possível confirmar que o AR esteja acoplado à lignina e desempenhando papel de subunidade não-canônica, por mais que as evidências possam sugerir este fato. Uma estratégia para verificar o acoplamento de AR à lignina poderia ser a avaliação deste composto usando o material da parede celular, o CWR, após hidrólise e dosagem de AR por HPLC. Para confirmar a presença do AR poderiam ser realizados estudos em ressonância magnética nuclear, para elucidar estruturalmente a composição da lignina das linhagens transgênicas.

Por fim, avaliando-se parâmetros de biomassa e inflorescências, para as duas construções não foram observadas diferenças estatisticamente significativas, sendo as plantas bem similares às selvagens em desenvolvimento e em aparência. Porém, seria interessante testar as linhagens para verificar o seu comportamento a partir de estresses bióticos e abióticos, com a finalidade de verificar a sua tolerância ou sensibilidade à fatores de estresses ambientais.

## 8. Conclusão

A partir dos testes das hipóteses deste trabalho é possível concluir que as plantas transgênicas, das duas construções, sofreram alterações na parede celular a partir da superexpressão de genes inseridos por engenharia genética. Foi possível verificar que a hipótese de que as plantas transgênicas poderiam liberar mais açúcar que plantas selvagens é verdadeira, além destas mesmas linhagens apresentarem modificações na composição e na quantidade de lignina presente na parede celular. Além disso, foi possível evidenciar que o metabolismo fenólico das plantas transgênicas foi alterado, sugerindo-se a hipótese de que as plantas se adaptaram para a estratégia envolvendo promotores constitutivos, e para a estratégia envolvendo promotores de parede celular secundária sustenta-se a hipótese de que o

compartilhamento dos precursores da via de AR e de lignina tem influência direta na composição monomérica da lignina destas plantas. Também foi possível confirmar que as plantas transgênicas não apresentaram características fenotípicas distintas de plantas selvagens.

Dessa forma, podemos afirmar a manipulação dos genes teve influência na composição da parede celular e influenciou positivamente a digestibilidade da biomassa lignocelulósica, contribuindo para uma hidrólise enzimática mais eficaz. Estes resultados são corroborados pela literatura e podem servir como base para estratégias de modificação de cana-de-açúcar, para o avanço da produção de etanol de segunda geração no Brasil e no mundo.

## 9. Referências

- ACKER, R.; VANHOLME, R.; PIENS, K.; BOERJAN, W. Saccharification Protocol for Small-scale Lignocellulosic Biomass Samples to Test Processing of Cellulose into Glucose. **BIO-PROTOCOL**, v. 6, n. 1, p. 1–9, 2016.
- AFEDZI, A. E. K.; PARAKULSUKSATID, P. Recent advances in process modifications of simultaneous saccharification and fermentation (SSF) of lignocellulosic biomass for bioethanol production. **Biocatalysis and Agricultural Biotechnology**, v. 54, n. October, p. 102961, 2023. Disponível em: <<https://doi.org/10.1016/j.bcab.2023.102961>>.
- ALARCÓN, A. L.; PALACIOS, L. M.; OSORIO, C.; NARVÁEZ, P. C.; HEREDIA, F. J.; ORJUELA, A.; HERNANZ, D. Chemical characteristics and colorimetric properties of non-centrifugal cane sugar (“panela”) obtained via different processing technologies, **Food Chemistry**, Volume 340, 2021, 128183, ISSN 0308-8146, <https://doi.org/10.1016/j.foodchem.2020.128183>.
- BEVILAQUA, J. M.; FINGER-TEIXEIRA, A.; MARCHIOSI, R.; DE OLIVEIRA, D. M.; JOIA, B. M.; FERRO, A. P.; PARIZOTTO, Â. V.; DOS SANTOS, W. D.; FERRARESE-FILHO, O. Exogenous application of rosmarinic acid improves saccharification without affecting growth and lignification of maize. **Plant Physiology and Biochemistry**, v. 142, n. June, p. 275–282, 2019.
- BEWG, W. P.; POOVAIAH, C.; LAN, W.; RALPH, J.; COLEMAN, H. D. RNAi downregulation of three key lignin genes in sugarcane improves glucose release without reduction in sugar production. **Biotechnology for Biofuels**, v. 9, n. 1, p. 1–13, 2016.
- BOERJAN, W.; RALPH, J.; BAUCHER, M. Lignin Biosynthesis. **Annual Review of Plant Biology**, v. 54, p. 519–546, 2003.
- BONDIA-PONS, I.; SAVOLAINEN, O.; TORRONEN, R.; MARTINEZ, J. A.; POUTANEN, K.; HANHINEVA, K. Metabolic profiling of Goji berry extracts for discrimination of geographical origin by non-targeted liquid chromatography coupled to quadrupole time-of-flight mass spectrometry. **Food Research International**. 2014; 63:132-138. doi: 10.1016/j.foodres.2014.01.067.
- BRASILEIRO, A. C. M.; CARNEIRO, V. T. de C. **Manual de Transformação Genética de Plantas**. Brasília, DF: EMBRAPA, 2015. v. 11543 p.

CESARINO, I.; DELLO IOIO, R.; KIRSCHNER, G. K.; OGDEN, M. S.; PICARD, K. L.; RAST-SOMSSICH, M. I.; SOMSSICH, M. Plant science's next top models. **Annals of Botany**, v. 126, n. 1, p. 1–23, 2020.

CHEN, F.; ZHUO, C.; XIAO, X.; PENDERGAST, T. H.; DEVOS, K. M. A rapid thioacidolysis method for biomass lignin composition and triclin analysis. **Biotechnol Biofuels** 14, 18 (2021). Disponível em: <<https://doi.org/10.1186/s13068-020-01865-y>>.

CHANDEL, A. K.; FORTE, M. B. S.; GONÇALVES, I. S.; MILESSI, T. S.; ARRUDA, P. V.; CARVALHO, W.; MUSSATTO, S. I. Brazilian biorefineries from second generation biomass: critical insights from industry and future perspectives. **Biofuels, Bioproducts and Biorefining**, v. 15, n. 4, p. 1190–1208, 2021.

CHRISTENSEN, C. S.L. ; RASMUSSEN, S. K. Low Lignin Mutants and Reduction of Lignin Content in Grasses for Increased Utilisation of Lignocellulose. **Agronomy**, 2019, 9, 256. Disponível em: <<https://doi.org/10.3390/agronomy9050256>>.

COOMEY, J. H.; SIBOUT, R.; HAZEN, S. P. Grass secondary cell walls, *Brachypodium distachyon* as a model for discovery. **New Phytologist**, 2020, 227(6), 1649-1667. <https://doi.org/10.1111/nph.16603>.

DA SILVA, A. R. G.; TORRES ORTEGA, C. E.; RONG, B. G. Techno-economic analysis of different pretreatment processes for lignocellulosic-based bioethanol production. **Bioresource Technology**, v. 218, p. 561–570, 2016.

DALENA, F.; SENATORE, A.; BASILE, M.; MARINO, D.; BASILE, A. **From sugars to ethanol-from agricultural wastes to algal sources: An overview**. [s.l.] Elsevier Inc., 2019. 3–34 p.

DE BUCK, S.; DE PAEPE, A.; DEPICKER, A. Transgene Expression in Plants , Control of. *In*: CHRISTOU, P.; SAVIN, R.; COSTA-PIERCE, B. A.; MISZTAL, I.; WHITELAW, C. B. A. **Sustainable food production**. New York, NY: Springer New York, 2013. p. 1570–1593.

DE C. L. E PENALVA SANTOS, D.; CORREA, C.; AMARAL ALVES, Y.; GOMES SOUZA, C.; A. MANCEBO BOLOY, R. Brazil and the world market in the development of technologies for the production of second-generation ethanol. **Alexandria Engineering Journal**, v. 67, p. 153–170, 2023. Disponível em: <<https://doi.org/10.1016/j.aej.2022.09.004>>.

OLIVEIRA, D. M.; FINGER-TEIXEIRA, A.; RODRIGUES MOTA, T.; SALVADOR, V. H.;

MOREIRA-VILAR, F. C.; CORREA MOLINARI, H. B.; CRAIG MITCHELL, R. A.; MARCHIOSI, R.; FERRARESE-FILHO, O.; DANTAS DOS SANTOS, W. Ferulic acid: A key component in grass lignocellulose recalcitrance to hydrolysis. **Plant Biotechnology Journal**, v. 13, n. 9, p. 1224–1232, 2015.

DE SOUZA, W. R.; MARTINS, P. K.; FREEMAN, J.; PELLNY, T. K.; MICHAELSON, L. V.; SAMPAIO, B. L.; VINECKY, F.; RIBEIRO, A. P.; DA CUNHA, B. A. D. B.; KOBAYASHI, A. K.; DE OLIVEIRA, P. A.; CAMPANHA, R. B.; PACHECO, T. F.; MARTARELLO, D. C. I.; MARCHIOSI, R.; FERRARESE-FILHO, O.; DOS SANTOS, W. D.; TRAMONTINA, R.; SQUINA, F. M.; CENTENO, D. C.; GASPAR, M.; BRAGA, M. R.; TINÉ, M. A. S.; RALPH, J.; MITCHELL, R. A. C.; MOLINARI, H. B. C. Suppression of a single BAHD gene in *Setaria viridis* causes large, stable decreases in cell wall feruloylation and increases biomass digestibility. **New Phytologist**, v. 218, n. 1, p. 81–93, 2018.

DE SOUZA, W. R.; PACHECO, T. F.; DUARTE, K. E.; SAMPAIO, B. L.; DE OLIVEIRA MOLINARI, P. A.; MARTINS, P. K.; SANTIAGO, T. R.; FORMIGHIERI, E. F.; VINECKY, F.; RIBEIRO, A. P.; DA CUNHA, B. A. D. B.; KOBAYASHI, A. K.; MITCHELL, R. A. C.; DE SOUSA RODRIGUES GAMBETTA, D.; MOLINARI, H. B. C. Silencing of a BAHD acyltransferase in sugarcane increases biomass digestibility. **Biotechnology for Biofuels**, v. 12, n. 1, p. 1–14, 2019. Disponível em: <<https://doi.org/10.1186/s13068-019-1450-7>>.

DEGREENIA, J.; WYNNE, G. Report Name : Biofuels Annual - Brazil. **Unites States Department of Agriculture Foreign Agriculture Service**, n. April 2020, p. 11, 2023. Disponível em: <<https://www.fas.usda.gov/data/malaysia-biofuels-annual-3>>.

ERMAWAR, R. A.; COLLINS, H. M.; BYRT, C. S.; HENDERSON, M.; O'DONOVAN, L. A.; SHIRLEY, N. J.; SCHWERDT, J. G.; LAHNSTEIN J., FINCHER, G. B.; BURTON, R. A. Genetics and physiology of cell wall polysaccharides in the model C4 grass, *Setaria viridis* spp. **BMC plant biology**, vol. 15 236. 2 Oct. 2015, doi:10.1186/s12870-015-0624-0.

EUDES, A.; DUTTA, T.; DENG, K.; JACQUET, N.; SINHA, A.; BENITES, V. T.; BAIDOO, E. E. K.; RICHE, A.; SATTLER, S. E.; NORTHEN, T. R.; SINGH, S.; SIMMONS, B. A.; LOQUÉ, D. SbCOMT (Bmr12) is involved in the biosynthesis of triclin-lignin in sorghum. **PLoS ONE**, v. 12, n. 6, p. 1–11, 2017.

FAN, L.; LINKER R.; GEPSTEIN S.; TANIMOTO E.; YAMAMOTO R.; NEUMANN P. M. Progressive inhibition by water deficit of cell wall extensibility and growth along the elongation

zone of maize roots is related to increased lignin metabolism and progressive stelar accumulation of wall phenolics. **Plant Physiol**, 2006, 140, 603–612. doi: 10.1104/pp.105.073130.

FOSTER, C. E.; MARTIN, T. M.; PAULY, M. Comprehensive compositional analysis of plant cell walls (Lignocellulosic biomass) part II: Carbohydrates. **Journal of Visualized Experiments**, n. 37, p. 10–13, 2010.

GALLEGO-GIRALDO, L.; POSÉ, S.; PATTATHIL, S.; PERALTA, A. G.; HAHN, M. G.; AYRE, B. G.; SUNUWAR, J.; HERNANDEZ, J.; PATEL, M.; SHAH, J.; RAO, X.; KNOX, J. P.; DIXON, R. A. Elicitors and defense gene induction in plants with altered lignin compositions. **New Phytologist**, v. 219, n. 4, p. 1235–1251, 2018.

GARCIA, T. C.; DURAND-MORAT, A.; YANG, W.; POPP, M.; SCHRECKHISE, W. Consumers' willingness to pay for second-generation ethanol in Brazil. **Energy Policy**, v. 161, p. 112729, 2022. Disponível em: <<https://doi.org/10.1016/j.enpol.2021.112729>>.

GARDINER, J. C.; TAYLOR, N. G.; TURNER, S. R. Control of cellulose synthase complex localization in developing xylem. **Plant Cell**, v. 15, n. 8, p. 1740–1748, 2003.

GOUVEIA-FIGUEIRA, S. C.; CASTILHO, P. C. Phenolic screening by HPLC–DAD–ESI/MSn and antioxidant capacity of leaves, flowers and berries of *Rubus grandifolius* Lowe, **Industrial Crops and Products**, Vol 73, 2015, p. 28-40, ISSN 0926-6690, <https://doi.org/10.1016/j.indcrop.2015.03.022>.

GRABBER, J. H.; DAVIDSON, C.; TOBIMATSU, Y.; KIM, H.; LU, F.; ZHU, Y.; OPIETNIK, M.; SANTORO, N.; FOSTER, C. E.; YUE, F.; RESS, D.; PAN, X.; RALPH, J. Structural features of alternative lignin monomers associated with improved digestibility of artificially lignified maize cell walls. **Plant Science**, v. 287, n. September 2018, p. 110070, 2019. Disponível em: <<https://doi.org/10.1016/j.plantsci.2019.02.004>>.

GRAYER, R. J.; KITE, G. C.; ABOU-ZAID, M.; ARCHER, L. J. The application of atmospheric pressure chemical ionisation liquid chromatography–mass spectrometry in the chemotaxonomic study of flavonoids: characterisation of flavonoids from *Ocimum gratissimum* var. *gratissimum*. **Phytochemical analysis**, Vol 11, Issue 4, July/August 2000, Pages 257-267.

GUAN, H.; LUO, W.; BAO, B.; CAO, Y.; CHENG, F.; YU, S.; FAN, Q.; ZHANG, L.; WU, Q.; SHAN, M. A Comprehensive Review of Rosmarinic Acid: From Phytochemistry to

Pharmacology and Its New Insight. **Molecules**, 2022, 27, 3292. Disponível em: <<https://doi.org/10.3390/molecules27103292>>.

HAAS, K. T.; PEAUCELLE, A. From monocots to dicots: The multifold aspect of cell wall expansion. **Journal of Experimental Botany**, v. 72, n. 5, p. 1511–1513, 2021.

HALPIN, C. Lignin engineering to improve saccharification and digestibility in grasses. **Current Opinion in Biotechnology**, v. 56, p. 223–229, 2019. Disponível em: <<https://doi.org/10.1016/j.copbio.2019.02.013>>.

HERNÁNDEZ-BLANCO, C.; FENG, D. X.; HU, J.; SÁNCHEZ-VALLET, A.; DESLANDES, L.; LLORENTE, F.; BERROCAL-LOBO, M.; KELLER, H.; BARLET, X.; SÁNCHEZ-RODRÍGUEZ, C.; ANDERSON, L. K.; SOMERVILLE, S.; MARCO, Y.; MOLINAA, A. Impairment of cellulose synthases required for Arabidopsis secondary cell wall formation enhances disease resistance. **Plant Cell**, v. 19, n. 3, p. 890–903, 2007.

HERNANDÉZ-LÓPEZ A.; DANIEL A.; FÉLIZ, S.; SIERRA, Z.; BRAVO, I. G.; DINKOVA, T. D.; AVILA-ALEJANDRE, A. Quantification of Reducing Sugars Based on the Qualitative Technique of Benedict. **ACS Omega**, 2020, 5 (50), 32403-32410, DOI: 10.1021/acsomega.0c04467.

HORTSCH, R.; CORVO, P. The Biorefinery Concept: Producing Cellulosic Ethanol from Agricultural Residues. **Chemie-Ingenieur-Technik**, v. 92, n. 11, p. 1803–1809, 2020.

HOU, X., SHAO, F., MA, Y.; LU, S. The phenylalanine ammonia-lyase gene family in *Salvia miltiorrhiza*: genome-wide characterization, molecular cloning and expression analysis. **Mol Biol Rep** 40, 4301–4310 (2013). <https://doi.org/10.1007/s11033-013-2517-3>.

HOENGENAERT, L.; WOUTERS, M.; KIM, H.; DE MEESTER, B.; MORREEL, K.; VANDERSYPPE, S.; POLLIER, J.; DESMET, S.; GOEMINNE, G.; RALPH, J.; BOERJAN, W.; VANHOLME, R. Overexpression of the scopoletin biosynthetic pathway enhances lignocellulosic biomass processing. **Science Advances**, 2022. v. 8, n. 28, 10.1126/sciadv.abo5738.

HUANG, J.; GU, M.; LAI, Z.; FAN, B.; SHI, K.; ZHOU, Y. H.; YU, J. Q.; CHEN, Z. Functional analysis of the Arabidopsis PAL gene family in plant growth, development, and response to environmental stress. **Plant Physiology**, v. 153, n. 4, p. 1526–1538, 2010.

JÖNSSON, L. J.; MARTÍN, C. Pretreatment of lignocellulose: Formation of inhibitory by-

products and strategies for minimizing their effects. **Bioresource Technology**, v. 199, p. 103–112, 2016.

JUNG, J. H.; ALTPETER, F. TALEN mediated targeted mutagenesis of the caffeic acid O-methyltransferase in highly polyploid sugarcane improves cell wall composition for production of bioethanol. **Plant Molecular Biology**, v. 92, n. 1–2, p. 131–142, 2016. Disponível em: <<http://dx.doi.org/10.1007/s11103-016-0499-y>>.

JIANG, Y. Q.; FANG, Z. X.; LEONARD, W.; ZHANG, P. Z. Phenolic compounds in Lycium berry: composition, health benefits and industrial applications. **Journal of Functional Foods**, 1756-4646, 77, 2021, DOI 10.1016/j.jff.2020.104340.

KARUNARATHNA, M. S.; SMITH, R.C. Valorization of Lignin as a Sustainable Component of Structural Materials and Composites: Advances from 2011 to 2019. **Sustainability**, 2020, 12, 734. Disponível em: <<https://doi.org/10.3390/su12020734>>.

KIM, J. S.; LEE, Y. Y.; KIM, T. H. A review on alkaline pretreatment technology for bioconversion of lignocellulosic biomass, **Bioresource Technology**, vol 199, 2016, p. 42-48, ISSN 0960-8524, <https://doi.org/10.1016/j.biortech.2015.08.085>.

KIM, Y. B.; KIM, J. K.; UDDIN, M. R.; XU, H.; PARK, W. T.; TUAN, P. A.; LI, X.; CHUNG, E.; LEE, J. H.; PARK, S. U. Metabolomics Analysis and Biosynthesis of Rosmarinic Acid in *Agastache rugosa* Kuntze Treated with Methyl Jasmonate. **PLoS ONE**, v. 8, n. 5, p. 1–8, 2013.

KISA, D.; ELMASTAS, M.; OZTURK, L.; KAYIR, O. Responses of the phenolic compounds of *Zea mays* under heavy metal stress. **Appl. Biol. Chem.** 2016, 59, 813–820. doi: 10.1007/s13765-016-0229-9.

KISHIMOTO, T.; URAKI, Y.; UBUKATA, M. Easy synthesis of  $\beta$ -O-4 type lignin related polymers. **Organic & Biomolecular Chemistry**, 2005, 1067 – 1073, v. 3, n. 6, <http://dx.doi.org/10.1039/B416699J>.

KORDALA, N.; WALTER, M.; BRZOZOWSKI, B.; LEWANDOWSKA, M. 2G-biofuel ethanol: an overview of crucial operations, advances and limitations. **Biomass Conversion and Biorefinery**, v. 14, n. 3, p. 2983–3006, 2024. Disponível em: <<https://doi.org/10.1007/s13399-022-02861-y>>.

LATA, C.; GUPTA, S.; PRASAD, M. Foxtail millet: A model crop for genetic and genomic studies in bioenergy grasses. **Critical Reviews in Biotechnology**, v. 33, n. 3, p. 328–343, 2013.

LEVSH, O.; PLUSKAL, T.; CARBALLO V.; MITCHELL A. J.; WENG J. K. Independent evolution of rosmarinic acid biosynthesis in two sister families under the Lamiids clade of flowering plants. **J Biol Chem.** 2019, Oct 18; 294(42): 15193-15205. doi: 10.1074/jbc.

LIU Q. Q.; ZHENG L.; HE F.; ZHAO F. J.; SHEN, Z. G.; ZHENG, L. Q. Transcriptional and physiological analyses identify a regulatory role for hydrogen peroxide in the lignin biosynthesis of copper-stressed rice roots. **Plant Soil**, 2015. V. 387, 323–336. doi: 10.1007/s11104-014-2290-7.

LIU, Q.; LUO, L.; ZHENG, L. Lignins: Biosynthesis and biological functions in plants. **International Journal of Molecular Sciences**, v. 19, n. 2, 2018.

LIU, T.; MIAO, P.; SHI, Y.; TANG, K. H. D.; YAP, P. S. Recent advances, current issues and future prospects of bioenergy production: A review. **Science of the Total Environment**, v. 810, p. 152181, 2022. Disponível em: <<https://doi.org/10.1016/j.scitotenv.2021.152181>>.

LOPES, M. L.; PAULILLO, S. C. de L.; GODOY, A.; CHERUBIN, R. A.; LORENZI, M. S.; GIOMETTI, F. H. C.; BERNARDINO, C. D.; DE AMORIM NETO, H. B.; DE AMORIM, H. V. Ethanol production in Brazil: a bridge between science and industry. **Brazilian Journal of Microbiology**, v. 47, p. 64–76, 2016. Disponível em: <<http://dx.doi.org/10.1016/j.bjm.2016.10.003>>.

LIU J.; WAN Y.; ZHAO Z.; CHEN, H. Determination of the content of rosmarinic acid by HPLC and analytical comparison of volatile constituents by GC-MS in different parts of *Perilla frutescens* (L.) **Britt. Chem Cent J.** 2013 Apr 1;7(1):61. doi: 10.1186/1752-153X-7-61.

LOQUÉ, D.; SCHELLER, H. V.; PAULY, M. Engineering of plant cell walls for enhanced biofuel production. **Current Opinion in Plant Biology**, v. 25, p. 151–161, 2015.

LU, F.; WANG, C.; CHEN, M., YUE, F.; RALPH, J. A facile spectroscopic method for measuring lignin content in lignocellulosic biomass. **Green Chemistry**, v. 23, p. 5106-5112, 2021. Disponível em: <<https://doi.org/10.1039/D1GC01507A>>.

MADEIRA F.; MADHUSOODANAN, N.; LEE J.; EUSEBI A.; NIEWIELSKA A.; TIVEY A. R. N.; LOPEZ, R.; BUTCHER S. The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. **Nucleic Acids Res.** 2024. Jul; 52(W1):W521-W525. doi: 10.1093/nar/gkae241. PMID: 38597606; PMCID: PMC11223882.

MAEDA, H. Lignin biosynthesis: Tyrosine shortcut in grasses. **Nature Plants** 2, 16080 (2016).

<https://doi.org/10.1038/nplants.2016.80>.

MANGA-ROBLES A.; SANTIAGO, R.; MALVAR, R. A.; MORENO-GONZALEZ, V.; FORNALE, S.; LOPEZ, I. Elucidating compositional factors of maize cell walls contributing to stalk strength and lodging resistance. **Plant Sci.** 2021. 307. doi: 10.1016/j.plantsci.2021.110882.

MARTIN, A.P.; PALMER, W.M.; BROWN, C. A developing *Setaria viridis* internode: an experimental system for the study of biomass generation in a C4 model species. **Biotechnol Biofuels** 9, 45 (2016). <https://doi.org/10.1186/s13068-016-0457-6>.

MARTINS, P. K.; CUNHA, B. A. D. B.; KOBAYSHI, A. K.; MOLINARI, H. B. C. *Setaria viridis* as a Model Plant for Functional Genomic Studies in C4 Crops. In: KUMAR, S.; BARONE, P.; SMITH, M. **Transgenic Plants: Methods and Protocols**. Methods in Molecular Biology. 1. ed. New York, NY: Springer New York, 2018. p. 49–65.

MARTINS, P. K.; MAFRA, V.; DE SOUZA, W. R.; RIBEIRO, A. P.; VINECKY, F.; BASSO, M. F.; DA CUNHA, B. A. D. B.; KOBAYASHI, A. K.; MOLINARI, H. B. C. Selection of reliable reference genes for RT-qPCR analysis during developmental stages and abiotic stress in *Setaria viridis*. **Scientific Reports**, v. 6, n. May, p. 1–10, 2016. Disponível em: <<http://dx.doi.org/10.1038/srep28348>>.

MARTINS, P. K.; RIBEIRO, A. P.; CUNHA, B. A. D. B. Da; KOBAYASHI, A. K.; MOLINARI, H. B. C. A simple and highly efficient *Agrobacterium*-mediated transformation protocol for *Setaria viridis*. **Biotechnology Reports**, v. 6, p. 41–44, 2015. Disponível em: <<http://dx.doi.org/10.1016/j.btre.2015.02.002>>.

MENG, X.; RAGAUSKAS, A. J. Recent advances in understanding the role of cellulose accessibility in enzymatic hydrolysis of lignocellulosic substrates. **Current Opinion in Biotechnology**, v. 27, p. 150–158, 2014. Disponível em: <<http://dx.doi.org/10.1016/j.copbio.2014.01.014>>.

MOREIRA-VILAR, F. C.; SIQUEIRA-SOARES, R. D. C.; FINGER-TEIXEIRA, A.; DE OLIVEIRA, D. M.; FERRO, A. P.; DA ROCHA, G. J.; FERRARESE, M. D. L. L.; DOS SANTOS, W. D.; FERRARESE-FILHO, O. The acetyl bromide method is faster, simpler and presents best recovery of lignin in different herbaceous tissues than klason and thioglycolic acid methods. **PLoS ONE**, v. 9, n. 10, 2014.

MOTTIAR, Y.; VANHOLME, R.; BOERJAN, W.; RALPH, J.; MANSFIELD, S. D. Designer

lignins: Harnessing the plasticity of lignification. **Current Opinion in Biotechnology**, v. 37, p. 190–200, 2016. Disponível em: <<http://dx.doi.org/10.1016/j.copbio.2015.10.009>>.

OLIVEIRA, D. M.; MOTA, T. R.; GRANDIS, A.; DE MORAIS, G. R.; DE LUCAS, R. C.; POLIZELI, M. L. T. M.; MARCHIOSI, R.; BUCKERIDGE, M. S.; FERRARESE-FILHO, O.; DOS SANTOS, W. D. Lignin plays a key role in determining biomass recalcitrance in forage grasses. **Renewable Energy**, v. 147, p. 2206–2217, 2020.

PERACCHI, L. M.; PANAHABADI, R.; BARROS-RIOS, J.; BARTLEY, L. E.; SANGUINET, K. A. Grass lignin: biosynthesis, biological roles, and industrial applications. **Frontiers in Plant Science**. v. 15, 2024. ISSN. 1664-462X, doi:10.3389/fpls.2024.1343097.

PETERSEN, M.; ABDULLAH, Y.; BENNER, J.; EBERLE, D.; GEHLEN, K.; HÜCHERIG, S.; JANIÁK, V.; KIM, K. H.; SANDER, M.; WEITZEL, C.; WOLTERS, S. Evolution of rosmarinic acid biosynthesis. **Phytochemistry**, v. 70, n. 15–16, p. 1663–1679, 2009. Disponível em: <<http://dx.doi.org/10.1016/j.phytochem.2009.05.010>>.

PINTO, L. da S. Fenilpropanoides da *Paspalum atratum*. 2015. 84 f. Dissertação (mestrado) - Universidade Estadual Paulista Júlio de Mesquita Filho, Instituto de Química., 2015.

PROTHMANN, J.; LI, K.; HULTEBERG, C.; SPÉGEL, P.; SANDAHL, M.; TURNER, C. Nontargeted Analysis Strategy for the Identification of Phenolic Compounds in Complex Technical Lignin Samples. **ChemSusChem**, 13(17), 2020, 4605–4612. <https://doi.org/10.1002/cssc.202000951>.

RALPH, J. Hydroxycinnamates in lignification. **Phytochemistry Reviews**, v. 9, n. 1, p. 65–83, 2010.

RALPH, J.; LAPIERRE, C.; BOERJAN, W. Lignin structure and its engineering, **Current Opinion in Biotechnology**, vol. 56, 2019, p. 240-249, ISSN 0958-1669, Disponível em: <<https://doi.org/10.1016/j.copbio.2019.02.019>>.

ROBAK, K.; BALCEREK, M. **Current state-of-the-art in ethanol production from lignocellulosic feedstocks** *Microbiological Research* Elsevier GmbH., 2020. Disponível em: <<https://doi.org/10.1016/j.micres.2020.126534>>.

SCHMITTGEN, T. D.; LIVAK, K. J. Analyzing real-time PCR data by the comparative CT method. **Nature Protocols**, v. 3, n. 6, p. 1101–1108, 2008.

SHAFIEI, R.; HOOPER, M.; MCCLELLAN, C.; OAKEY, H.; STEPHENS, J.; LAPIERRE,

C.; TSUJI, Y.; GOEMINNE, G.; VANHOLME, R.; BOERJAN, W.; RALPH, J.; HALPIN, C. Downregulation of barley ferulate 5-hydroxylase dramatically alters straw lignin structure without impact on mechanical properties. **Frontiers in Plant Science**, v. 13, n. January, p. 1–15, 2023.

SILVEIRA, M. R. S. da.; OSTER, A. H.; MOURA, C. F. H.; SILVA, E. de O.; SILVA, L. M. A.; SOUSA, A. E. D. de; Protocolos para avaliação das características físicas e físico-químicas, dos compostos bioativos e atividade antioxidante do pedúnculo do caju. **Embrapa Agroindústria Tropical**. 43 p. 2018.

SOLTANIAN, S.; AGHBASHLO, M.; ALMASI, F.; HOSSEINZADEH-BANDBAFHA, H.; NIZAMI, A. S.; OK, Y. S.; LAM, S. S.; TABATABAEI, M. A critical review of the effects of pretreatment methods on the exergetic aspects of lignocellulosic biofuels. **Energy Conversion and Management**, v. 212, n. February, p. 112792, 2020. Disponível em: <<https://doi.org/10.1016/j.enconman.2020.112792>>.

SOMERVILLE, C. Cellulose synthesis in higher plants. **Annual Review of Cell and Developmental Biology**, v. 22, p. 53–78, 2006.

STUDER, M. H.; DEMARTINI, J. D.; DAVIS, M. F.; SYKES, R. W.; DAVISON, B.; KELLER, M.; TUSKAN, G. A.; WYMAN, C. E. Lignin content in natural *Populus* variants affects sugar release. **Proc Natl Acad Sci U S A**. 2011;108(15):6300–5.

SUN, S.; SUN, S.; CAO, X.; SUN, R. The role of pretreatment in improving the enzymatic hydrolysis of lignocellulosic materials. **Bioresource Technology**, v. 199, p. 49–58, 2016. Disponível em: <<http://dx.doi.org/10.1016/j.biortech.2015.08.061>>.

TAN, L.; SUN, W.; LI, X.; ZHAO, J.; QU, Y.; CHOO, Y. M.; LOH, S. K. Bisulfite pretreatment changes the structure and properties of oil palm empty fruit bunch to improve enzymatic hydrolysis and bioethanol production. **Biotechnol J**. 2015; 10(6):915–25.

TOBIMATSU, Y.; ELUMALAI, S.; GRABBER, J. H.; DAVIDSON, C. L.; PAN, X.; RALPH, J. Hydroxycinnamate conjugates as potential monolignol replacements: In vitro lignification and cell wall studies with rosmarinic acid. **ChemSusChem**, v. 5, n. 4, p. 676–686, 2012.

TRÓCSÁNYI, E.; GYÖRGY, Z.; ZÁMBORINÉ-NÉMETH, É. New insights into rosmarinic acid biosynthesis based on molecular studies. **Current Plant Biology**, v. 23, n. January, p. 100162, 2020. Disponível em: <<https://doi.org/10.1016/j.cpb.2020.100162>>.

VANHOLME, R.; CESARINO, I.; RATAJ, K.; XIAO, Y.; SUNDIN, L.; GOEMINNE, G.; KIM, H.; CROSS, J.; MORREEL, K.; ARAUJO, P.; WELSH, L.; HAUSTRAETE, J.; MCCLELLAN, C.; VANHOLME, B.; RALPH, J.; SIMPSON, G. G.; HALPIN, C.; BOERJAN, W. Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in arabidopsis. **Science**, v. 341, n. 6150, p. 1103–1106, 2013.

VANHOLME, R.; DE MEESTER, B.; RALPH, J.; BOERJAN, W. Lignin biosynthesis and its integration into metabolism. **Current Opinion in Biotechnology**, v. 56, p. 230–239, 2019. Disponível em: <<https://doi.org/10.1016/j.copbio.2019.02.018>>.

VANHOLME, R.; DEMEDTS, B.; MORREEL, K.; RALPH, J.; BOERJAN, W. Lignin biosynthesis and structure. **Plant Physiology**, v. 153, n. 3, p. 895–905, 2010.

VEGA-SÁNCHEZ, M. E.; RONALD, P. C. Genetic and biotechnological approaches for biofuel crop improvement. **Current Opinion in Biotechnology**, v. 21, n. 2, p. 218–224, 2010. Disponível em: <<http://dx.doi.org/10.1016/j.copbio.2010.02.002>>.

VERMAAS, J. V.; CROWLEY, M. F.; BECKHAM, G. T. A Quantitative Molecular Atlas for Interactions Between Lignin and Cellulose. **ACS Sustainable Chemistry & Engineering**, 2019, 7 (24), 19570-19583, DOI: 10.1021/acssuschemeng.9b04648

VOINICIUC, C.; PAULY, M.; USADEL, B. Monitoring polysaccharide dynamics in the plant cell wall. **Plant Physiology**, v. 176, n. 4, p. 2590–2600, 2018.

WAGNER, A.; DONALDSON, L.; KIM, H.; PHILLIPS, L.; FLINT, H.; STEWARD, D.; TORR, K.; KOCH, G.; SCHMITT, U.; RALPH, J. Suppression of 4-coumarate-CoA ligase in the coniferous gymnosperm *Pinus radiata*. **Plant Physiology**, v. 149, n. 1, p. 370–383, 2009.

WEITZEL, C.; PETERSEN, M. Cloning and characterisation of rosmarinic acid synthase from *Melissa officinalis* L. **Phytochemistry**, v. 72, n. 7, p. 572–578, 2011. Disponível em: <<http://dx.doi.org/10.1016/j.phytochem.2011.01.039>>.

XIAO, Y.; ZHANG, L.; GAO, S.; SAECHAO, S.; DI, P.; CHEN, J.; CHEN, W. The *c4h*, *tat*, *hpr* and *hpd* genes prompted engineering of rosmarinic acid biosynthetic pathway in *Salvia miltiorrhiza* hairy root cultures. **PLoS ONE**, v. 6, n. 12, 2011.

ZHANG, J. H.; LY, H. Z.; LIU, W. J.; JI A. J.; ZHANG, X.; SONG, J. Y.; LUO, H. M.; CHEN, S. L. bHLH transcription factor SmbHLH92 negatively regulates biosynthesis of phenolic acids and tanshinones in *Salvia miltiorrhiza*. **Chin Herb Med**. 2020 Apr 19;12(3):237-246. doi:

10.1016/j.chmed.2020.04.001. PMID: 36119017; PMCID: PMC9476745.

ZHANG, B.; GAO, Y.; ZHANG, L.; ZHOU, Y. The plant cell wall: Biosynthesis, construction, and functions. **Journal of Integrative Plant Biology**, v. 63, n. 1, p. 251–272, 2021.

ZHANG, L.; ZHOU, Y.; ZHANG, B. Xylan-directed cell wall assembly in grasses. **Plant Physiology**, v. 194, n. 4, p. 2197–2207, 2024. Disponível em: <<https://doi.org/10.1093/plphys/kiad665>>.

ZHAO, X. Q.; GUO, S.; YAN, H.; LU, Y. Y.; ZHANG, F. QIAN, D.W.; WANG, H. Q.; DUAN, J. A. Analysis of phenolic acids and flavonoids in leaves of *Lycium barbarum* from different habitats by ultra-high-performance liquid chromatography coupled with triple quadrupole tandem mass spectrometry. **Biomed Chromatogr.** 2019 Aug;33(8):e4552. doi: 10.1002/bmc.4552. Epub 2019 Jun 3. PMID: 30985939.

YAN, C.; LI C.; JIANG, M.; XU, Y.; ZHANG, S.; HU, X.; CHEN, Y.; LU, S. Systematic characterization of gene families and functional analysis of PvRAS3 and PvRAS4 involved in rosmarinic acid biosynthesis in *Prunella vulgaris*. **Frontiers in Plant Science**, 15, 2024, 1664-462X, DOI=10.3389/fpls.2024.1374912.

YAN, G.; KAISER, K. A rapid and sensitive method for the analysis of lignin phenols in environmental samples using ultra-high performance liquid chromatography-electrospray ionization-tandem mass spectrometry with multiple reaction monitoring, **Analytica Chimica Acta**, Volume 1023, 2018, Pages 74-80, ISSN 0003-2670, <https://doi.org/10.1016/j.aca.2018.03.054>.

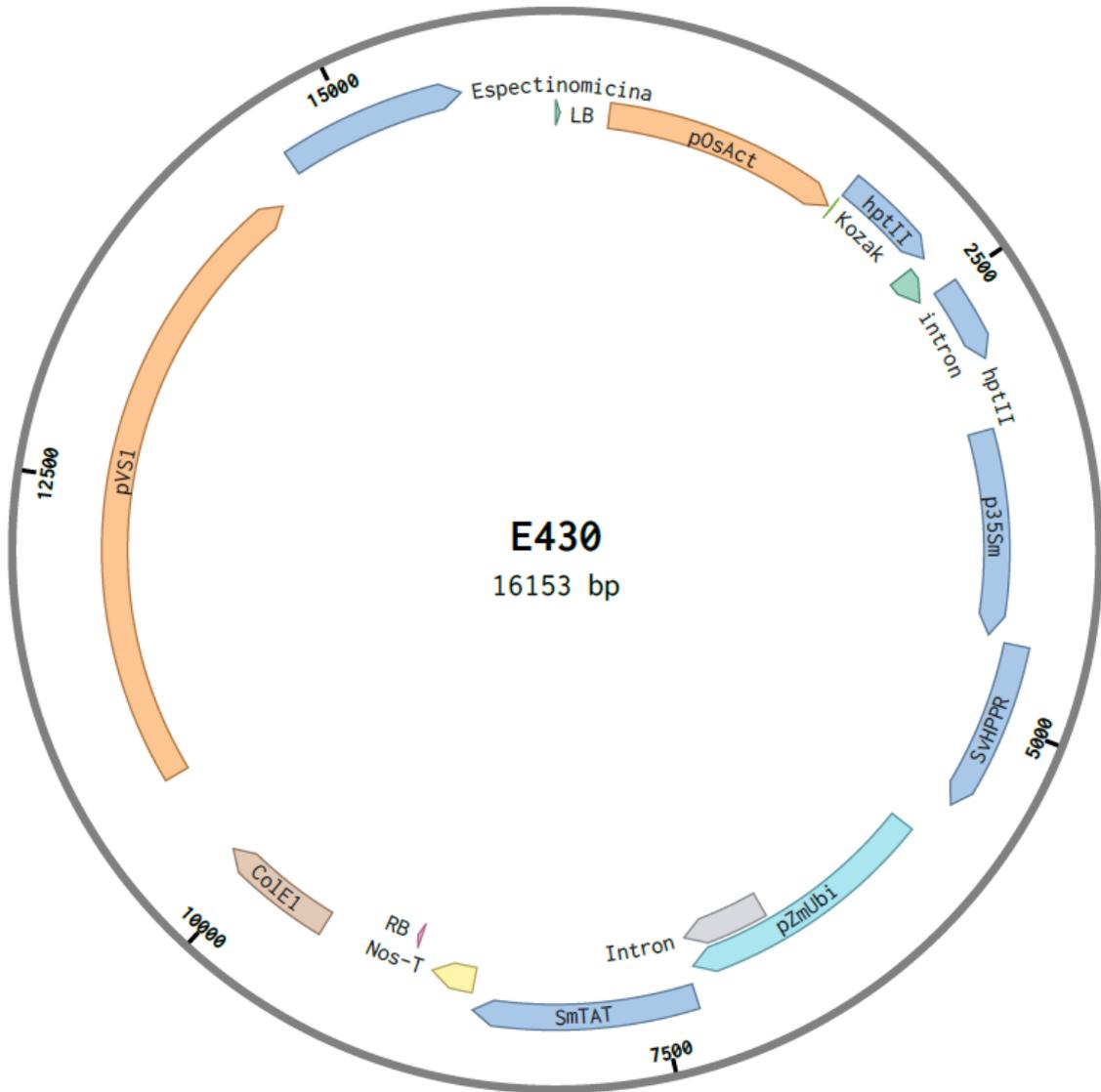
YANG, B.; WYMAN, C. E. BSA Treatment to enhance enzymatic hydrolysis of cellulose in lignincontaining substrates, **Biotechnol. Bioeng.**, 94 (4) (2006), pp. 611-617.

YAMAMOTO, F.Y.; PÉREZ-LOPEZ, C.; LOPEZ-ANTIA, A. Linking MS1 and MS2 signals in positive and negative modes of LC-HRMS in untargeted metabolomics using the ROIMCR approach. **Anal Bioanal Chem** 415, 6213–6225 (2023). <https://doi.org/10.1007/s00216-023-04893-3>.

YE, M.; YANG, W. Z.; LIU, K. D.; QIAO, X.; LI, B. J.; CHENG, J.; FENG, J.; GUO, D. A.; ZHAO, Y. Y. Characterization of flavonoids in *Millettia nitida* var. *hirsutissima* by HPLC/DAD/ESI-MS<sub>n</sub>, **Journal of Pharmaceutical Analysis**, Volume 2, Issue 1, 2012, Pages 35-42, ISSN 2095-1779, <https://doi.org/10.1016/j.jpha.2011.09.009>.

## Anexo I – Representação dos vetores e sequências

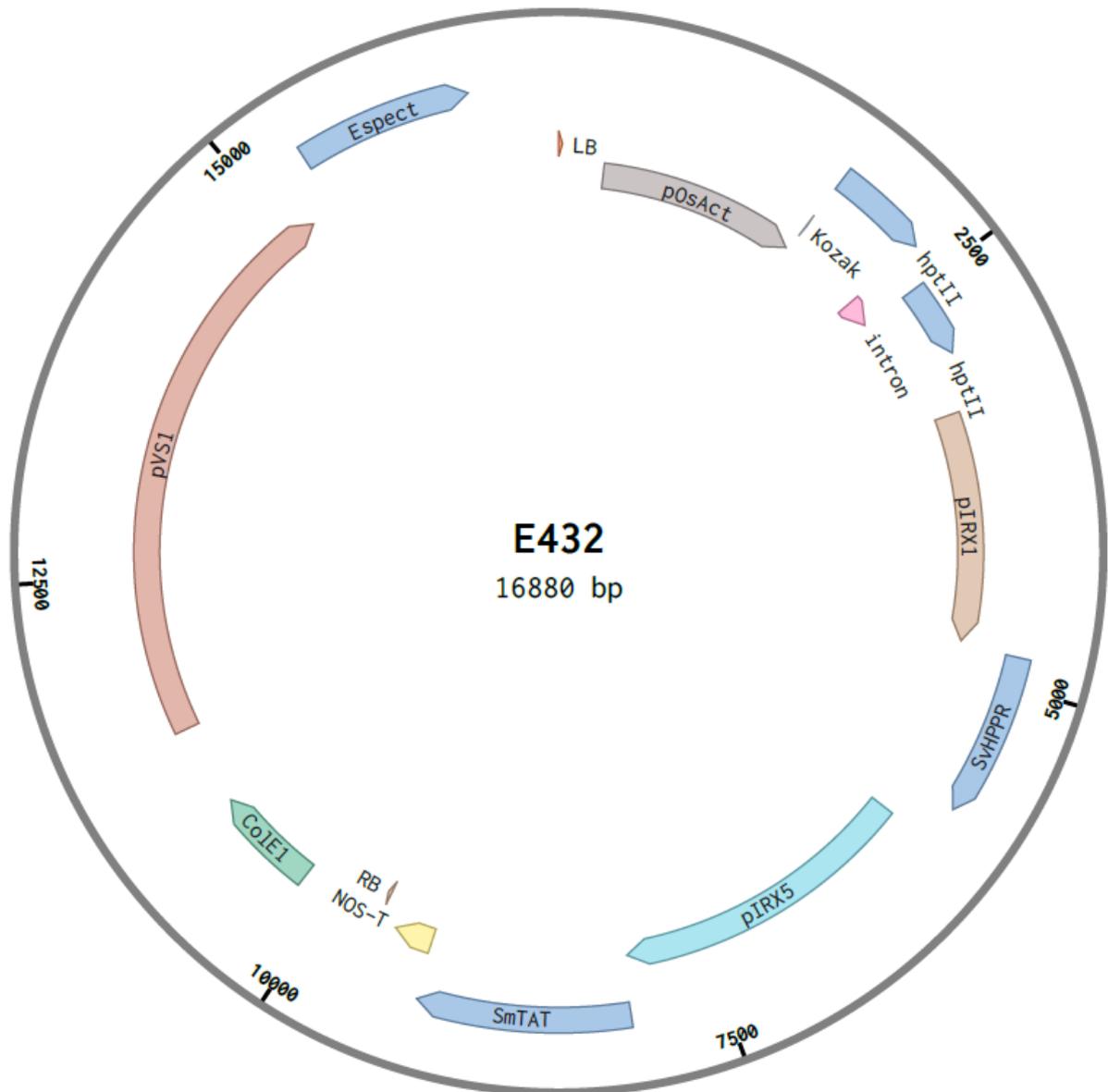
**Figura 1** – Representação esquemática do vetor E430.



Nota: *Espectinomicina*, gene de resistência ao antibiótico; *LB*, *left border* ou borda esquerda; *pOsAct*, promotor; *Kozak*, sinal traducional sítio Kozak; *hptII*, gene de resistência à Higromicina, dois sítios e um íntron; *p35Sm*, promotor; *SvHPPR*, gene; *pZmUbi*, promotor; *SmTAT*, gene; *Nos-T*, terminador; *RB*, *right border* ou borda direita; *ColE1*, origem de replicação *E. coli*; *pVSI*, origem de replicação para *A. tumefaciens*.

**Fonte:** do autor (2024).

**Figura 2** – Representação esquemática do vetor E432.



Nota: *Espectinomicina*, gene de resistência ao antibiótico; *LB*, *left border* ou borda esquerda; *pOsAct*, promotor; *Kozak*, sinal traducional sítio Kozak; *hptII*, gene de resistência à Higromicina, dois sítios e um íntron; *pIRX1*, promotor; *SvHPPR*, gene; *pIRX5* promotor; *SmTAT*, gene; *Nos-T*, terminador; *RB*, *right border* ou borda direita; *ColE1*, origem de replicação *E. coli*; *pVSI*, origem de replicação para *A. tumefaciens*.

**Fonte:** do autor (2024).

**pOsAct** – Sequência otimizada de CDS promotor de actina em *Oryza sativa* otimizado para plantas monocotiledôneas, presente na construção como promotor do gene de resistência à Higromicina. Presente nas duas construções, E430 e E432.

```
tcgaggtcattcatatgcttgagaagagagtcgggatagtcacaaataaaacaaggaattacacctggtcaaaagtgaacacatcagttaaaaggtgtataaagt
aaaatcggtaataaaagtgcccaagtgaattactcttttactattataaaaattgaggatgttttgcgtactttgatacgtcattttgtatgaattgtttta
agtttattcgtttggaatgcatatctgtattgagtcgggttfaagtcgtttgttaatacagaggattgtataaagaatatttaaaaaacccatagccta
atttgacataattttgagaaaaatataattcaggcgaattctcaacaataaagaattaaaatagctttccccctgtcagcgcagtggtatttttctagtaaaa
ataaaagataaaacttagactcaaacatttcaaaaacaacccctaaagtctcaaaagcccaagtgctatccacgatccatagcaagccagcccaaccaacca
accaaccacccagtcagccaactggacaatagctccacacccccactatcaccgtgagttgtccgcacgcaccgcacgtctcgcagccaaaaaa
aaagaagaaaaaaagaaaaagaaaaacagcaggtgggtccggctgfgggggccgaaacgcgaggagatcgcgagccagcagcagggccggccc
tccctcgtctcaaaagaaacgccccctcgcactatatacataccccctctctctccatcccccaacctaccaccaccaccaccacctccacctct
ccccctcgtcgcggacgagcagctctccccctccccctcgcgcgcggcgggtaaccacccgccccctctctcttttctctcgtttttttctcgtct
cgtctcgtatcttgcttggtgatttgggtggcgagagcggctctgctgcgcccagatcggctgcggggagggcgggatctcgcggtgggctctcgc
ccggcgtggatccggcccggatctcgggggaatgggctctcggatgtagatctcgcgacccgctgttgggggagatgatgggggttaaaattccgcat
gtaaaacaagatcaggaaggggaaaaagggcactatggtttatattttatatttctgctctcgcagccttagatgtgctagatcttcttcttctttgtggtag
aatttgaatccctcagcattgtcatcgttagtttttcttcatgattgtgacaaatgcagcctcgtcggagctttttgtagta
```

**hptII** – Sequência otimizada de CDS gene de resistência ao antibiótico Higromicina, *hygromycin phosphotransferase II*. Presente nas duas construções, E430 e E432.

```
ccgaaatgaagaaaccggagctgaccgccactccgtcgaaggtcctcagagaagttgatagcgtgtctgacctatgcagttgtcggagggcgaagagtc
ccgcgctgacgtttgatgttgccggcgccgggtacgtcctccgcgtgaactcttgcgtgacggctctataaagatcggtagctttacagccatttcgctcggc
ggcctcgtctatcccgaagtcctcgcacatcggcgagttttccgagagccttacgtattgtattcccggcgctcaaggggtgacctccaggtatcggcga
ctgagttgctcgtggttccagcccgtcggcggcgtatggacccatcggcgccggatcttctcaaacgtcgggctcggcgctcggccctcaggggat
tgccagtagaccacttgaggagctttatctgcgctatcggcagcgcagctgtagatcattggcaaacggtatggacgatacgtctccgcgagcgtggcccag
ctctggacgaactcatgttggggcggaggtgccccgaggtaaagtctctctacctttgatataataaataaattatcattatagtagtaataatattcaata
tttttcaaaataaaagaatgtagtatatagcaattgctttctgtagttataagtgtagtatattataacttttctaataatagacaaaattgttgatgtagcgtcc
gccacctggtcatcggactcggcttaataacgtctcaccgataatggcggtactcctcgtcagctggtccgaagctatgtttggcgattcgcagtagc
aagtggcgaacattttcttctgaggccttgcttgcctgcatggaacagcaaacgcgctactttgagcggaggcaccggagctggcggggtctccccgctcag
ggcctatattgtgagcgtcggcctcggaccgctttaccagtcctggtgatgggaattcgcagatgctgctggcgcaaggccgctgtagcgtattgtcagga
gcgcgccggggaccgtggggcgcactcagatcggcgcggtcggctcgggttgacggatgggtgcgtcgaagtgctcggacagcggcaacagggcg
cctccaccgccccgggccaaggagtgta
```

**hptII** – Sequência de aminoácidos

```
MKKPELTATSVEKFLIEKFDVS DLMQLSEGEESRAFSFDVGGRGYVLRVNSCADGFYKDRYVYRHFA
SAALPIPEVLDIGEFSESLTYCISRR AQVTLQDLPETELPAVLQPVAEAMDAIAAADLSQTSFGPFGPQ
GIGQYTTWRDFICAIADPHVYHWQTV MDDTVSASVAQALDELMLWAEDCPEVSFCFYL*YIYNNYH*L
VVI*YFKYFFQNKRM*YIAIAFL*FISVYILY NFNISNI*PKFVDVQVRHLVHADFGSNNVLT DNGRITAVID
WSEAMFGDSQYEVANIFFWRPWLACME QQTRYFERRHPELAGSPRLRAYMLRIGLDQLYQSLVDGNF
DDAAWAQGRCDIVRSGAGTVGRTQIARRSAAVWTDGCVEVLADSGNRRPSTRPRAKE*
```

**p35S** – Sequência otimizada de CDS promotor *CaMV 35S* modificado para otimização de expressão em monocotiledôneas. Utilizado para superexpressão do gene *SvHPPR* na construção E430.

```
ggtccgattgagactttcaacaaaggtaatatccggaacctcctcggattccattgccagctatctgctactttattgtgaagatagtggaaggaaggtggct
cctacaatgccatcattgcgataaaggaaagccatcgtgaagatgcctctgccagcagtggtcccaaatggacccccaccacagggagcagcgtgga
aagaagacgtccaaccagctctcaagcaagtggattgatgtgatgctccgattgagactttcaacaaaggtaatatccggaacctcctcggattccattgcc
```

agctatctgtcactttattgtgaagatagtgaaaaggaaggtggctcctacaaatgccatcattgcgataaaggaaaggccatcgttgaagatgctctgccgacag  
 tggcccaaatgatgacccccaccacagggagcatcgtggaagaagacgttcaaccacgcttcaaagaagtgattgatgatctccactgacgta  
 agggatgacgcacaatcccactatccttcgaagaccctctctataaggaagttcattcatttgagaggacacgctgacaagctgactctagcagatcctcta  
 gaacctctccacacactcaagccacactattggagaacacacagggacaacacaccataagatccaagggaggcctccgccgcccggtaaccacccgc  
 ccctctcctctttctcctggttttttccgtctcgtctcgtatcttggccttgtagttgggtggcgagaggcggcttcgtgcgcccagatcgggtcgcggga  
 ggggcggtatcgcggctgggctcgcggcgtggatccggccggatcgcggggaatgggctcctggatgatagatcgcgacccgctgttgggg  
 gagatgatgggggttaaaattccgccgtgtaacaagatcaggaagaggggaaaaggcactatggttatatttatatttctgctcctcgcagcgttag  
 atgtgctagatcttcttcttcttgggtagaattgaatccctcagcattgtcatcggtagttttcttctcatgattgtgacaatgcagcctcgtcggagctttt  
 gtagtagaag

**SvHPPR** – Sequência CDS gene *HPPR hidroxiphenylpyruvate reductase* em *S. viridis*.

Identificador Phytozome Sevir.5G009200. Utilizado para as duas construções, E430 e E432.

Atggagtcctcgggtctctcttccacccatgaacgcttacttgagcaggaactcgtcgtcgtcagcgtcatcgttctggaggaccacccggcc  
 cgcctcgggacgagttcctcagggccacgctggaagcatccgtgccgtggtgggaatgcttctatggtcggatgccctctattgacgccctcccgcctc  
 gagatcgtcgcattctccgtggcattgatagggtagctcgtcaagtgcagagaactggtatcagggtcaccaactcctgatgttggaccgacgatgc  
 gccgacttagctgtggactcgcctatagcggctcgtcgcacgcatccacaggctgctgtttgtagggcggcttgaaggcaaggcgcactacgctctca  
 ctaccggttccagcgggaaggggtcgggaattatcggctcggccgatcggctcggcgtggcgtaaaagggttagaggcattcattgccccgtaactactatcaa  
 cggacaaagcaggactacctaattacagctattacccagtggtgcgagctagctccaactccgatccttgttggcttggccactcaaccccagaccgt  
 cacattgtcaacaggaagtgatggacgcgctgggacctaagggtgcttcatcaatattggcgcggacccacgttgatgagccggagctgctcgtccctg  
 gtggaggccgctcgtggtggcgtggtcttgcgcttcgaggatgagcctaacttccgaagcactctcactatggacaacgtggtcgtggtcccacatcgg  
 gagcggaaaccaatgagactaggaaggcatggccgatctggtctcggtaactggaggctcacgtgctcaagaagcctcttgcaccctgtcctctga

**SvHPPR** – Sequência de aminoácidos

MESLGVLLLHPMNAYLEQELDRRCRLHRFWEAPPGPPRDEFRLAHAGSIRAVVGNASYGADAALIDAL  
 PALEIVASFSVIGIDRVDLAKCRERIRVTNTPDVLDDVADLAVGLAIAVLRRIQADRFRVAGLWKAK  
 GDYALTRFSGKRVGIIGLGRIGLAVAKRVEAFDCPVNYYQRTKQDYPNYTYPSVVELAANSDLLVV  
 ACPLNAQTRHIVNREVMDALGPKGVLINIGRPHVDEPELVAALVEGRLGGAGLDVFEDEPNVPEALFT  
 MDNVVVPHIGSGTNETRKGMDLVLGNLEAHVLLKPLLPVL\*

**pZmUbi** – Sequência otimizada de CDS promotor constitutivo de Ubiquitina em *Zea mays*.

Utilizado para superexpressão do gene *SmTAT* na construção E430.

ccctgcttaaatgagatagcagacgcctatgatcgcgatgatttcttcaattctgttgcacgttgaacaaacctgagcatgtgtagctcagatcctaccgccg  
 gtttcggtcattctaataatatacaccggtactatcgtattttatgaataatctccgttcaattfactgattgctcgcagcgggtacagtgagcgttgactg  
 agtgcagcgtgaccggctgctccctctctagagataatgacattgcatgcttaagtataaaaaattaccacatattttttgacactgtttgaagtgcagttat  
 ctatcttatacatatattaaacttactctcagaataatatactatagtactacaataatcagtggttttagagaatcataaatgaacagtagacatggtctaaagg  
 acaattgagattttgacaacaggactctacagtttatcttttagtgcagtggtctccttttttgcacaaatagcttccatataataacttccatttttagtacatc  
 catttagggtttagggtaagtgttttatagactaatttttagtacatctattttattctatttttagcctctaaattaagaaaactaaaactctatttttagtttttataataatt  
 agatataaaatagaataaaatgaagtgactaaaataaacaataacccttaagaaaataaaaaactaagaaacattttctgtttcagtagataatgccagcctg  
 taaacgccgtcgcagcaggtcaaacggacaccaaccagcgaaccagcagcgtcgcgctggccaagcgaagcagacggcagcgcagcgtcctcgtcgtcctc  
 tggaccctctcagagattccgctccaccgttgacttgcctcgtcggcatccagaattgctggtggcggagcggcagcgtgagccggcagcggcagcggc  
 ctctctcctctcagcgcaccggcagctacgggggattccttcccaccgctcctcgttccctcctcgcgccgtaataaatagacacccccccacaccctc  
 ttcccaacctcgtgttgcggagcgcacacacacaaccagatctccccaaatccaccgctcggcactccgcttaaggtacgccgctcgtcctcccccc  
 cccccctctactctctagatcgcgcttccggtccatggttagggccggtagtctactctgttcatggttggtagatccgtttgtgttagatccgtcgtcgttag  
 cgttctacacggatgogactgtacgtcagacagcttctgattgctaacttccagtggttcttggggaatcctgggatgctcgttagccgttccgcagacgggat  
 cgtactagataggtatcatggtgatgtgggtttactgatcatatagatggcatatgcagcatctattcatatgcttaaccttgagtacctatctataataaac

aagatggtttataattttgatcttgatatacttggatgatggcatatgcagcagctatattggatttttagccctgccttcatacgtattttgcttgactgtttct  
ttgtcagatctcacctgtgtgtgttacttctgcaggt

**SmTAT** – Sequência otimizada de CDS gene *TAT Tyrosine aminotransferase* em *Salvia miltiorrhiza*. Baseado na sequência NCBI GenBank DQ334606.1. Utilizado para as duas construções, E430 e E432.

atggagctgcagaaccctgctcaggagatcgcaccaccactattaccatcaagggtattctcggccttctcatgtccagcactgacccaaggagcttgaaa  
gagggatcatgattgggatcggatcctaccgctacagctgtccacgcttcaatgccgcagggagggcgtggtgaggtctccgtagcacaaggctc  
aacggttatgctcccaccgctgactgcccgaactcgcgagggcattgctgagctctctcgtgaccttcttacaagctgcccgcagattcagctatgtacc  
gctgggtgcacgcagccatcgaattgcgctcctcgttggtagccagggcgaacatcctcctgctcgcacctgcttccatctacggtctctgcgtagc  
ttccgaacattgaagtgcggtactttgacctcaccggagcaggggtgggaagtcgactcgcagccgttcagatctagctgaccacaactgctgccatggtg  
atcattaaccccgaaaccatgcccgaacgtctacagttatcagcactgaagaaatcgtgagaccgcaagaggctcggcatagtagtgatgccgatgaa  
gtctacggctatcttcttggagcaaatcttctgctccatgggtatttggctcgcgctcctgtggtcactcctgtagcctgtccaagagatggtggcca  
gggtgctcctcggatgctcgtcattaacaccccgatggttctctgatgagccctaaagtcgtgagaggatcaagaactgtagatctgcccgggtccggc  
cacctcattcaggctgcccgcagatcgtgagcagacacaagaagcttcttccgtaagaccattaacatccttaagcagacttccgatatctgatatcagaag  
attgaggacatcaacgggattacctgcccactaagcgaaggagctatggcggttatggtgaaactcaatctgagcaggatgaaggatcagtgacgatcga  
cttctgctcaagctcgcgaaggagctgtcattatcttggccctcgtgtgttcttaagaactggattcgcacaccttccagtgatgctccctgcccctg  
aggaggtatggagcgtctcaagctcttctcgcgagggcacagctactga

**SmTAT** – Sequência de aminoácidos

MELQNPAQEIDAPTTITIKILGLLMSSTDPKESGKRVISLGIGDPTAYSCFHASNAAQEGVVEALRSTKF  
NGYAPTAGLPQTREAIAEYLSRDLPLYKLPADSVYVTAGCTQAIEIALSVLARPGANILLPRPCFPYGLCA  
SFRNIEVRYFDLHPEQGWVVDLDAVADLADHNTVAMVIINPGNPGNVYSYQHLKKAIAETAKRLGIVVI  
ADEVYGHAFGANPFVPMGIFGSIAPVVTLGSLSKRWLVPGWRLGWLVINDPDGSLMSPKFVERIKKY  
CDICGGPATFIQAAVPEIVEQTQE VFRKTINILKQTS DIC YQKIEDINGITCPTKPKGAMAFMVKLNLSR  
MKDISDDIDFCFLAKEESVIILPGLAVGLKNWIRITFAVDVPALEEAMERLKSFCERHSY\*

**pIRX1** – Sequência otimizada de CDS promotor *IRX1*. 1.5 kb à jusante do gene *IRX1* de *Setaria italica*; Identificador Phytozome Seita 5G319100. Utilizado para expressão espaço-temporal do gene *SvHPPR* na construção E432.

attagctgtacatgggtcgcgcccttggatcgggtgtaatactatgttccaagttctcacaaaacaagagaaaattcctatacggcactagataaaaacatgattcc  
ctatagactagaatgttgcctccttatttgatcgggatttattttcattccttctgacactttgtccttattgacactgttcttattgacactaggattatcttcc  
attccttcttacttccattagatcccttagaaacttccatcaaatcactcacaatgtaccaaaattctctgaaatttctactctcatgttgacgattctttgagc  
tgcccagtgcccatggtattcaggctgcccgcagggctgatgatcctcaatggcaagtggggtgaaagcgggtgggataatgatcaaggaaaggatgaaaa  
ataaattctaattgcaataaggaatataaagaaggaatgaaaaataatcctagtgtcaataagaggcaaaacttctagtgtcaggtagggaaccagatttcat  
gtgatccataatggtgaatttctaaaaaacaatcgcacggagggcagggcttgggtgacgcacatcaactagagcaaccacgcgacccagcggcgggac  
ctgattggctgattgtaacgcagcagcattgcgagaacacgcgtactacagtactagcaagttactgacgctgaacttgaagtccacggcatcaact  
cccccaagacgaagaagcaaatgccgacacgtgaaaaatgaccgatcctcttccaagggcagagcagacaatcctcactcgcgtcccgcctgctg  
gtcggcgtcgcgacttgggaacgggctgtcagaaaaacgacgccccagcgcgatctcccactgatgactccagtactcctcttcttccgaaccggcgacc  
caccaccgcaccaacggaaccaccaaccaaccgccccgcgacctcgcaccccccgttctcagggttacgccccgcggcgggaacgggagccaat  
cagtgaatcaccgctgcttctgcttgaacccccgaaaacctcgcggagaagaaccactgctgcacggcgcagcgaaggcagcagcagcacacacct  
ctcctcctcctctatggcctcaccacaccttcttcttcttccatcccaccagcctgctgcagcggcactccccactcaccaccgctgcccagcgc  
cacctcaccgaccgcccgcctcgtcacccccgcggcggatgggtgctgctgctatataccacgctccaacctccggcagcttagcttcttctgctct  
ccccattctcccaccgggtccccctggtgctgcccgcgcagcggatcccaaccagggtgctgctgaccgcggtgaggaggaggaggacc

**pIRX5** – Sequência otimizada de CDS promotor IRX5 1980 pb à jusante do start códon do gene IRX5 de *Arabidopsis thaliana* Identificador Phytozome At5g44030. Utilizado para expressão espaço-temporal do gene *SmTAT* na construção E432.

gaagccatcctctacctcgaaaaactgttgcgagaagaagacatgcgatggcatggatgcttgatctttgacattgatgacactcttctcaaccattccttacca  
caagagcaacgggtgttcgggtaataaaactaaacttaaccatatacattagccttgattcggttttggttgattatggatattaaagatccgaattatattgaacaaa  
aaaaaatgattatgtcacataaaaaaaattggcttgaatttggtttagatgggttaaatgtctacacctaatcattcatttctggttagcttaattcggtttagaat  
gaaaccgggattgacatgttacattgattgaaacagtggtgagcaactgaacacgaccaagttcgaggaaatggcaaaatcgggcaaggcaccagcggttccac  
acatggtgaagttgaccatgagatcagagagagaggttcaagatcttttgatctctctcgtaaagagtatctcagatctgccaccgtcgaatcttattgaagccg  
gttaccacagctggtctaaccctcttctgaggttcgaatcataatfaataaccgcattaaacgaaatttaattctaatccaccaaatcaaaaagtaaaactagaacactt  
cagataaatttctgcttctgtgacttcatttattcttaaacacaagaactatagaccataatcgaataaaaaccctaaaaccaaatttatctattfaaaacaacatt  
agctattgagttcttttaggtaagttattaaaggtttggagacttfaagatgtttcagcatttatggttgtgcattaatgttttagttagtaaagaagaaaagatagtaa  
ttaaagagttggttgaatcataatfaaaacattaataggtatttatgtctaatttggggacaaaatagtggaattctttatcatactagctagtcttatcagtttgaact  
cgggttatgattatgtacatgcaatggtcataataatctatgagcaatataaattcgagcattttggtataacataatgagccaagtatacaaaaagtatcaaaccta  
tcagggggagaagatgatgaaaagaagagtgtagccaatacaaaagcagatttgaggacatggccttacaagcttgggtacagagtttggggagtgatgggtgca  
caatggaacagcttctctggtgtccagttccaagagaacctcaagctccctaaactccatctactatgtcgcctgattaatcttattactaacaacaaataagatca  
gagttcattctgattctgagcttttttctctcctctttcatttctggtttataaacaatcaaatgcttatgatccatgcatgaacctgatcatctttgtgtttttttc  
cttctgtattaccatttgggctttgtgaaattgatttgggctttgttatataatctcctcttctcttctctacctgattggattcaagaacatagccagatttggtaaagttt  
ataagatacaaaaatattaagtaagactaaagtagaatacataataactgaaagctactctaaagtatacaaaattcaagaactcaaaaagaatacaaacagtagaa  
gttgaagctcaagcaatfaaattataaaaaactaactacactgagctgtctcctctccaccaaatctgttctgtctcttgaagcttcttatgacacaaaccttag  
acccaattcactcacagttgttacaacctcagtttctcacaacaaactcaaacatctacccttatattacctctttatctctcaatcatcaaaacacatagtcacatac  
atttctctacccacctctgctctgcttccgagagctcagtgacctgcc

## Anexo II - Tabelas

**Tabela 1** – Meio de cultura de indução de calos CIM:

Reagente	Concentração final	Para 100 mL	Para 500 mL
Sais MS (Sigma)	4,3 g/L	430 mg	2,15 g
Vitaminas Setaria/CIM 1000X*	1x	100 µL	500 µL
D-Biotina [0,5 mg/mL]**	1 mg/L	200 µL	1 mL
Sacarose	30 g/L	3 g	15 g
Inositol	100 mg/L	10 mg	50 mg
Fitigel	4 g/L	200 mg	2 g
Solução CuSO <sub>4</sub> .5H <sub>2</sub> O 1000X***	0,92 mg/L	100 µL	500 µL
Solução de 2,4-D [1 mg/mL]****	2 mg/L	200 µL	1 mL
Kinetina (1mg/mL)#	0,5 mg/L	50 µL	250 µL
Ajustar pH 5,8 com NaOH 1 M e autoclavar			

**Fonte:** (MARTINS *et al.*, 2015).

**Tabela 2** – Meio de co-inoculação CIM líquido:

Reagente	Concentração final	Para 100 mL	Para 500 mL
Sais MS (Sigma)	4,3 g/L	430 mg	2,15 g
Vitaminas Setaria/CIM 1000X	1x	100 µL	500 µL
D-Biotina 500X [0,5 mg/mL]	1 mg/L	200 µL	1 mL
Sacarose	30 g/L	3 g	15 g
Inositol	100 mg/L	10 mg	50 mg
Solução de 2,4-D [1 mg/mL]	2 mg/L	200 µL	1 mL
Kinetina (1 mg/mL)	0,5 mg/L	50 µL	250 µL
Acetoseríngona (100 mM)	200 mM	200 µL	1 mL
Ajustar pH 5,8 com NaOH 1 M e autoclavar			

**Fonte:** (MARTINS *et al.*, 2015).

**Tabela 3** – Meio de regeneração MRS:

Reagente	Concentração final	Para 100 mL	Para 500 mL
Sais MS (Sigma)	4,3 g/L	430 mg	2,15 g
Vitaminas Setaria/CIM2 1000X	1x	100 µL	500 µL
D-Biotina [0,5 mg/mL]	1 mg/L	200 µL	1 mL
Sacarose	20 g/L	2 g	10 g
Inositol	100 mg/L	10 mg	50 mg
Kinetina (1mg/mL)	2 mg/L	200 µL	1 mL
Fitagel	2 g/L	200 mg	1 g
Ajustar pH 5,8 com NaOH 1 M e autoclavar			
Timentin [300 mg/mL]	150 mg/L	50 µL	250 µL
Higromicina [50 mg/mL]	30 mg/L	60 µL	300 µL

**Fonte:** (MARTINS *et al.*, 2015).

**Tabela 4** – Meio de desenvolvimento MD:

Reagente	Concentração final	Para 100 mL	Para 500 mL
Sais MS (Sigma)	2,15 g/L	215 mg	1,07 g
Vitaminas MS 1000X	1x	100 µL	500 µL
Sacarose	15 g/L	1,5 g	7,5 g
Inositol	100 mg/L	10 mg	50 mg
Fitagel	2 g/L	200 mg	1 g
Ajustar pH 5,8 com NaOH 1M e autoclavar			
Timentin [300 mg/mL]	150 mg/L	50 µL	250 µL
Higromicina [50 mg/mL]	30 mg/L	60 µL	300 µL

**Fonte:** (MARTINS *et al.*, 2015).

**Tabela 5** – Sequências dos primers utilizados para confirmação das plantas por PCR:

PRIMER	SEQUÊNCIA	DESCRIÇÃO
HptII-F	GATGTTGGCGACCTCGTATT	Primers para amplificar sequência do gene de resistência à Higromicina
HptII-R	TGTTTATCGGCACTTTGCAT	
CAC-F	CTGCTTCTGGTCTTCGTGTT	Primers para amplificar um gene constitutivo de <i>S. viridis</i> utilizado para controle
CAC-R	CATCACGAGAGCAGGATCATAC	
pACT-F	CCGGGAAATCTACATGGATCAG	Primers para amplificar o promotor do gene de resistência à Higromicina
pACT-R	TTTGTCTGGTACTTTGATACGTCA	

**Fonte:** Do autor (2024).

**Tabela 6** – Meio de cultura ½ MS:

Reagente	Concentração final	100 mL	200 mL	500 mL	1000 mL
Sais MS	2,15 g/L	215 mg	430 mg	1,075 g	2,15 g
Vitaminas MS	1x	100 µL	200 µL	500 µL	1 mL
Inositol	100 mg/L	10 mg	20 mg	50 mg	100 mg
Fitigel	4 g/L	400 mg	800 mg	2 g	4 g
Ajustar pH 5,8 com NaOH 1M e autoclavar					

**Fonte:** (MARTINS *et al.*, 2015).

**Tabela 7** – Sequências dos primers utilizados nos experimentos de qPCR:

PRIMER	SEQUÊNCIA	DESCRIÇÃO
RT Svhprr-F-A	ATTCTCCGTGGGCATTGATAG	Primers para amplificar sequência do gene <i>SvHPPR</i> em plantas transgênicas
RT Svhprr-R-A	ATCGTCGGTCAACACATCAG	
RT Svhprr-F-B	CTGGAGATCGTCGCATCATT	Primers para amplificar sequência do gene <i>SvHPPR</i> em plantas transgênicas
RT Svhprr-R-B	GAGTGTTGGTGACCCTGATAC	
RT Smtat-F-A	CGATGCACCCACCACTATTA	Primers para amplificar sequência do gene <i>SmTAT</i> em plantas transgênicas
RT Smtat-R-A	CCCAAATGATGACCCTCTT	
RT Smtat-F-B	GCTTTCGGAGCAAATCCTTTC	Primers para amplificar sequência do gene <i>SmTAT</i> em plantas transgênicas
RT Smtat-R-B	CTGGCACCAACCATCTCTT	
RT Svhprr WT-F-1	CGAACGAAGCAGGACTACC	Primers para amplificar sequência do gene <i>SvHPPR</i> em plantas WT
RT Svhprr WT-R-1	CCCAGGGCATCCATCAC	
RT-Svhprr WT-F-2	AAGAGTGGGCATCATTGGG	Primers para amplificar sequência do gene <i>SvHPPR</i> em plantas WT
RT-Svhprr WT-R-2	CTTCGTTGCTGGTAGTAGTT	
RT CAC-F	CTGCTTCTGGTCTTCGTGTT	Primers para amplificar a sequência do gene <i>Clathrin adaptor complex</i> Sevir.1G284400
RT CAC-R	GTATGATCCTGCTCTCGTGATG	
RT SUI-F	CCAGAGCTTGGACAGGTCATTC	Primers para amplificar a sequência do gene <i>Translation fator SUI1</i> Sevir.2G348300
RT SUI-R	ACAATGCCAGCCTGGACAA	

**Fonte:** Do autor (2024).

**Tabela 9** – Etapas das rodadas de transformação genética de *Setaria viridis*.

	Calos	Transf.	Vetor	CIM + Aceto	CIM + Tim	CIM + Tim + Higro	MRS + Tim + Higro (1)	MRS + Tim + Higro (2)	Meio MD ndução de raízes e parte aérea		Aclimação
1	22/abr	30/abr	E430	30/abr	03/mai	10/mai	19/mai	09/jun	N/C		
2	30/abr	06/mai	E432	06/mai	11/mai	18/mai	26/mai	Placas contaminadas, foram descartadas			
3	14/mai	21/mai	E430	21/mai	25/mai	01/jun	08/jun	28/jun	N/C		
4	21/mai	28/mai	E432	28/mai	31/mai	07/jun	15/jun	05/jul	3 explantes em 12/abr	N/C	
5	03/jun	12/jun	E430	12/jun	15/jun	22/jun	30/jun	20/jul	1 candidato em 13/jul		1 ev. em 16/jul
6	10/jun	18/jun	E432	18/jun	21/jun	30/jun	07/jul	25/jul	vários explantes em 15/jul	3 explantes em 23/jul	2 ev. em 27/jul
7	24/jun	02/jul	E430	02/jul	05/jul	13/jul	20/jul	09/ago	3 explantes em 29/ago	N/C	
8	30/jun	08/jul	E432	09/jul	12/jul	19/jul	26/jul	23/ago	N/C		
9	02/jul	09/jul		09/jul	12/jul	19/jul	26/jul	23/ago	N/C		
10	16/jul	23/jul	E430	23/jul	27/jul	03/ago	10/ago	01/set	5 explante em 21/set	N/C	
11	03/ago	10/ago	E432	10/ago	13/ago	Placas contaminadas em Cim + Tim, foram descartadas					
12	25/ago	01/set	E430	01/set	04/set	10/set	17/set	08/out	vários explantes 28/out	3 explantes em 07/out	2 ev. em 15/out
13	31/ago	07/set	E432	07/set	10/set	17/set	24/set	14/out	2 explantes em 04/nov	N/C	
14	13/set	21/set	E430	21/set	24/set	01/out	08/out	28/out	vários explantes em 18/nov	N/C	
15	01/out	08/out	E432	08/out	11/out	18/out	25/out	15/nov	4 explante em 05/dez	N/C	
16	14/out	22/out	E430	Placas contaminadas em Cim + Aceto, foram descartadas							
17	18/out	26/out	E432	26/nov	29/out	05/nov	12/nov	02/dez	3 explantes em 22/dez	N/C	
18	25/out	05/nov	E430	05/nov	08/nov	15/nov	22/nov	12/dez	vários explante em 02/jan	N/C	
19	04/nov	12/nov	E432	12/nov	15/nov	22/nov	29/nov	20/dez	vários explantes em 10/jan	4 explantes em 17/dez	2 ev. em 20/dez
20	10/nov	19/nov	E430	19/nov	22/nov	29/nov	06/dez	26/dez	2 explantes em 20/jan	N/C	
21	18/nov	26/nov	E432	26/nov	29/nov	06/dez	13/dez	03/jan	3 explantes em 23/jan	N/C	
22	01/mar	08/mar	E430	08/mar	11/mar	18/mar	25/mar	15/abr	4 explantes em 12/abr	N/C	
23	21/mar	29/mar	E432	29/mar	01/abr	08/abr	08/abr	23/abr	N/C		
24	07/abr	15/abr	E432	15/abr	18/abr	25/abr	02/mai	22/mai	N/C		
25	11/abr	19/abr	E430	19/abr	22/abr	29/abr	06/mai	26/mai	Vários explantes em 06/jun	N/C	
26	18/abr	29/abr	E432	29/abr	02/mai	09/mai	Calos com contaminação constante independente da troca de meio				
27	25/abr	05/mai	E432	05/mai	09/mai	16/mai	23/mai	14/jun	N/C		
28	03/mai	12/mai	E430	12/mai	16/mai	23/mai	30/mai	20/jun	vários explantes em 21/jun	2 explantes em 29/jun	1 ev. em 05/jul
29	10/mai	19/mai	E430	19/mai	23/mai	30/mai	06/jun	26/jun	vários explantes em 04/jul	2 explantes em 06/jul	2 ev. em 11/jul
30	18/mai	27/mai	E430	27/mai	30/mai	06/jun	13/jun	03/jul			
31	31/mai	10/jun	E432	10/jun	13/jun	20/jun	27/jun	13/jul	4 explantes em 20/jul	1 explante em 27/jul	1 ev. em 30/jul
32	13/jun	24/jun	E430	24/jun	27/jun	Calos com contaminação constante independente da troca de meio					
33	21/jun	01/jul	E432	01/jul	04/jul	11/jul	18/jul	08/ago	vários explantes em 15/ago	2 explantes em 21/ago	2 ev. em 27/ago
34	29/jun	08/jul	E430	08/jul	11/jul	18/jul	25/jul	19/ago	3 explantes em 21/ago	N/C	
35	12/jul	22/jul	E432	22/jul	25/jul	01/ago	08/ago	28/ago	N/C		
36	20/jul	29/jul	E432	29/jul	01/ago	08/ago	15/ago	05/set	N/C		
37	01/ago	09/ago	E430	09/ago	12/ago	19/ago	26/ago	16/set	N/C		

<b>38</b>	10/ago	19/ago	<b>E430</b>	19/ago	22/ago	29/ago	05/set	25/set	N/C	
<b>39</b>	23/ago	02/set	<b>E432</b>	02/set	05/set	12/set	19/set	09/out	vários explantes em 12/out	1 ev. em 17/out
<b>40</b>	06/set	15/set	<b>E432</b>	15/set	18/set	25/set	03/out	23/out	3 explantes em 24/out	2 ev. em 29/out
<b>41</b>	22/nov	22/nov	<b>E430</b>	01/dez	04/dez	11/dez	18/dez	05/jan	vários explantes em 07/jan	N/C
<b>42</b>	29/nov	29/nov	<b>E430</b>	08/dez	11/dez	18/dez	23/dez	05/jan	3 explantes em 01/jan	1 ev. em 10/jan
<b>43</b>	08/dez	08/dez	<b>E430</b>	15/dez	18/dez	23/dez	03/jan	13/jan	2 explantes em 15/jan	1 ev. em 22/jan

N/C: Nenhum candidato. Aclimação refere-se à transferência de um explante candidato para o solo. Transformações 1 a 21 foram realizadas durante o ano de 2021; transformações 22 a 43 foram realizadas em 2022.

**Fonte:** Do autor (2024).

**Tabela 10** – Avaliação por PCR de plantas da linhagem E430.

	Identificação da planta										
	1	2	3	4	5	6	7	8	9	10	
Linhagens E430	1	Sementes homozigotas									
	2	Sementes homozigotas									
	3	Sementes homozigotas									
	4	Positiva	N/C	N/C	N/C	N/C	N/C	Positiva	N/C	Positiva	Negativa
	5	Positiva	Positiva	Positiva	Positiva	Positiva	Positiva	Positiva	Negativa	Positiva	Positiva
	6	Positiva	Negativa	N/C	Positiva	Positiva	Negativa	Positiva	Positiva	Positiva	Positiva
	7	N/C	N/C	N/C	N/C	N/C	N/C	N/C	Positiva	N/C	N/C
	8	Positiva	Positiva	Positiva	Negativa	Positiva	Positiva	Positiva	Positiva	Positiva	Positiva

**Nota:** PCR's utilizando primers *pACT* de todas as linhagens da construção E430. Positiva refere-se àquela planta cuja PCR resultou em uma amplificação, considerando-se então uma planta positiva; Negativa refere-se àquela planta cuja PCR não resultou em uma amplificação, considerando-se então uma planta negativa; N/C refere-se ao não crescimento. As linhagens 1, 2 e 3, cujas sementes já eram homozigotas, não foram testadas para PCR. As sementes T1 da linhagem 8 foram consideradas homozigotas para os experimentos de crescimento e biomassa.

**Tabela 11** – Avaliação por PCR de plantas da linhagem E432.

	Identificação da planta										
	1	2	3	4	5	6	7	8	9	10	
Linhagens E430	1	Sementes homozigotas									
	2	Negativa	Negativa	Negativa	Negativa	Negativa	Negativa	Negativa	Negativa	Negativa	Negativa
	3	Positiva	Positiva	Positiva	Positiva	Positiva	Positiva	Positiva	Positiva	Positiva	Negativa
	4	Positiva	Positiva	Positiva	Negativa	Positiva	Positiva	Positiva	Positiva	Positiva	Positiva
	5	Positiva	Positiva	Positiva	Positiva	Positiva	Positiva	Positiva	Positiva	Positiva	Positiva
	6	Positiva	Positiva	N/C	Positiva	Positiva	Positiva	Positiva	N/C	Positiva	Positiva
	7	N/C	N/C	Positiva	Negativa	Positiva	Negativa	Negativa	Negativa	N/C	Negativa
	8	Negativa	Negativa	Negativa	Negativa	Positiva	Negativa	Negativa	Positiva	Positiva	Negativa
	9	Negativa	Negativa	Negativa	Positiva						
	10	Positiva	Negativa	Negativa	Positiva	Negativa	Negativa	Negativa	Positiva	Positiva	Positiva

**Nota:** PCR's utilizando primers *pACT* de todas as linhagens da construção E432. Positiva refere-se àquela planta cuja PCR resultou em uma amplificação, considerando-se então uma planta positiva; Negativa refere-se àquela planta cuja PCR não resultou em uma amplificação, considerando-se então uma planta negativa; N/C refere-se ao não crescimento. As plantas da linhagem 1, cujas sementes já eram homozigotas, não foram testadas para PCR. As sementes T1 das linhagens 3, 4 e 5 foram consideradas homozigotas para os experimentos de crescimento e biomassa.

**Tabela 12** – Divisão de plantas para composição dos pools das linhagens de E430.

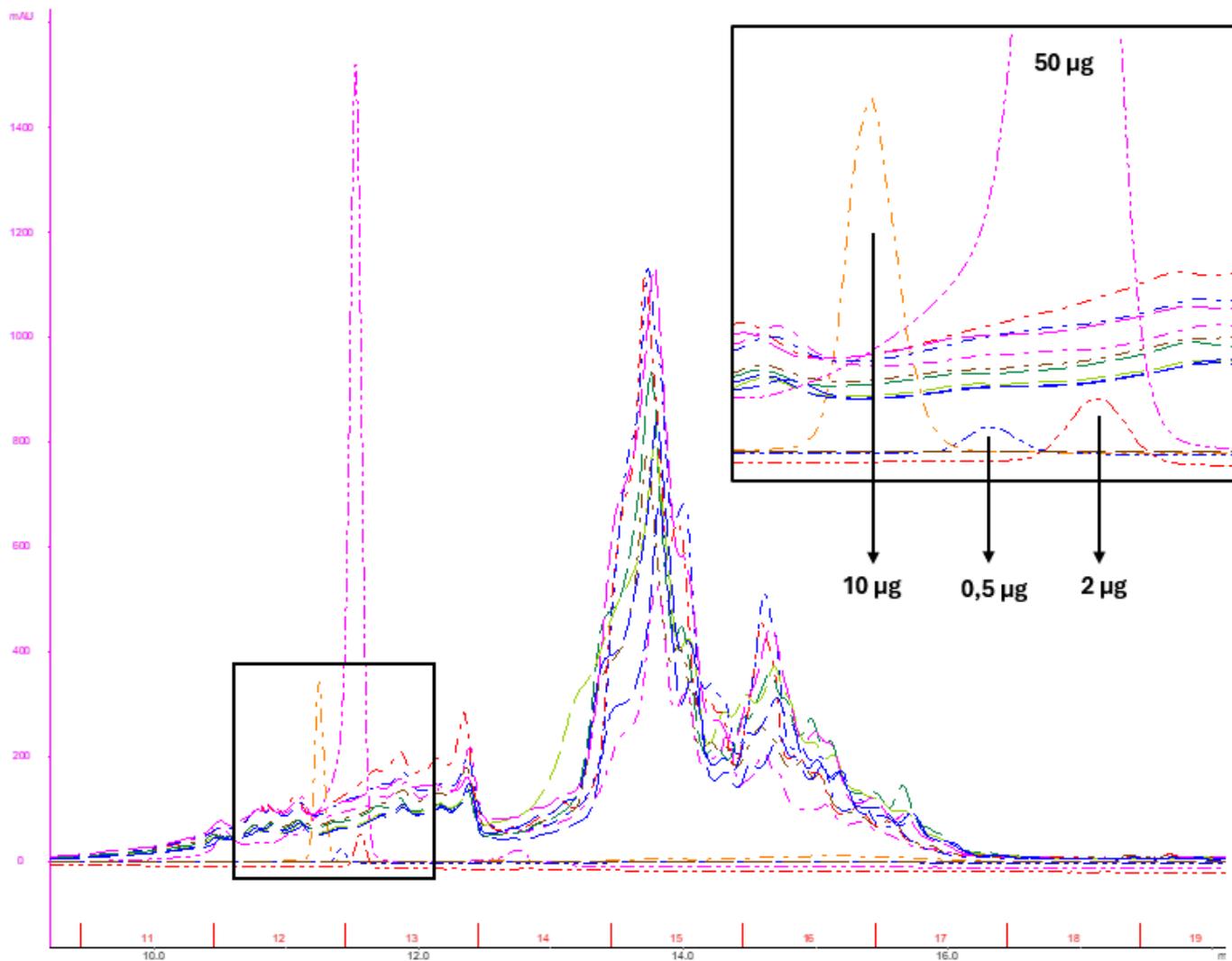
<b>LINHAGEM</b>	<b>POOL</b>	<b>PLANTAS</b>	<b>TOTAL DE PLANTAS</b>
<b>Linhagem 1</b>	E430 Ev1 - A	2, 3, 4	9
	E430 Ev1 - B	1, 8, 9	
	E430 Ev1 - C	5, 10, 7	
<b>Linhagem 2</b>	E430 Ev2 - A	2, 3, 5	9
	E430 Ev2 - B	1, 4, 7	
	E430 Ev2 - C	6, 8, 9	
<b>Linhagem 3</b>	E430 Ev3 - A	5, 7, 1	9
	E430 Ev3 - B	4, 6, 10	
	E430 Ev3 - C	2, 3, 8	
<b>Linhagem 4</b>	E430 Ev4 - A	7	3
	E430 Ev4 - B	9	
	E430 Ev4 - C	1	
<b>Linhagem 5</b>	E430 Ev5 - A	3, 7, 6	8
	E430 Ev5 - B	10, 4	
	E430 Ev5 - C	1, 2, 9	
<b>Linhagem 6</b>	E430 Ev6 - A	1, 4, 10	7
	E430 Ev6 - B	5, 7	
	E430 Ev6 - C	8, 9	
<b>Linhagem 7</b>	E430 Ev7 - A	8	1
<b>Linhagem 8</b>	E430 Ev8 - A	3, 6, 9	9
	E430 Ev8 - B	5, 7, 10	
	E430 Ev8 - C	1, 2, 8	
<b>Plantas selvagens (WT)</b>	E430 WT - A	1, 2, 3	9
	E430 WT - B	10, 7, 8	
	E430 WT - C	5, 6, 9	

**Tabela 13** – Divisão de plantas para composição dos pools das linhagens E432.

<b>LINHAGEM</b>	<b>POOL</b>	<b>PLANTAS</b>	<b>TOTAL DE PLANTAS</b>
<b>Linhagem 1</b>	E432 Ev1 - A	3, 4, 5	9
	E432 Ev1 - B	6, 7, 10	
	E432 Ev1 - C	1, 2, 8	
<b>Linhagem 3</b>	E432 Ev3 - A	2, 3, 5	9
	E432 Ev3 - B	1, 6, 8	
	E432 Ev3 - C	4, 7, 9	
<b>Linhagem 4</b>	E432 Ev4 - A	3, 7, 10	9
	E432 Ev4 - B	2, 8, 9	
	E432 Ev4 - C	6, 5, 1	
<b>Linhagem 5</b>	E432 Ev5 - A	4, 7, 9	9
	E432 Ev5 - B	2, 3, 10	
	E432 Ev5 - C	5, 6, 8	
<b>Linhagem 6</b>	E432 Ev6 - A	10, 4, 7	8
	E432 Ev6 - B	5, 6, 9	
	E432 Ev6 - C	1, 2	
<b>Linhagem 7</b>	E432 Ev7 - A	3	2
	E432 Ev7 - B	5	
<b>Linhagem 8</b>	E432 Ev8 - A	5	3
	E432 Ev8 - B	9	
	E432 Ev8 - C	8	
<b>Linhagem 9</b>	E432 Ev9 - A	10, 8	7
	E432 Ev9 - B	5, 6	
	E432 Ev9 - C	4, 7, 9	
<b>Linhagem 10</b>	E432 Ev10 - A	10, 9	5
	E432 Ev10 - B	1, 8	
	E432 Ev10 - C	4	
<b>Plantas selvagens (WT)</b>	E432 WT - A	2, 3, 9	9
	E432 WT - B	1, 7, 10	
	E432 WT - C	4, 8, 5	

### Anexo III – Cromatogramas

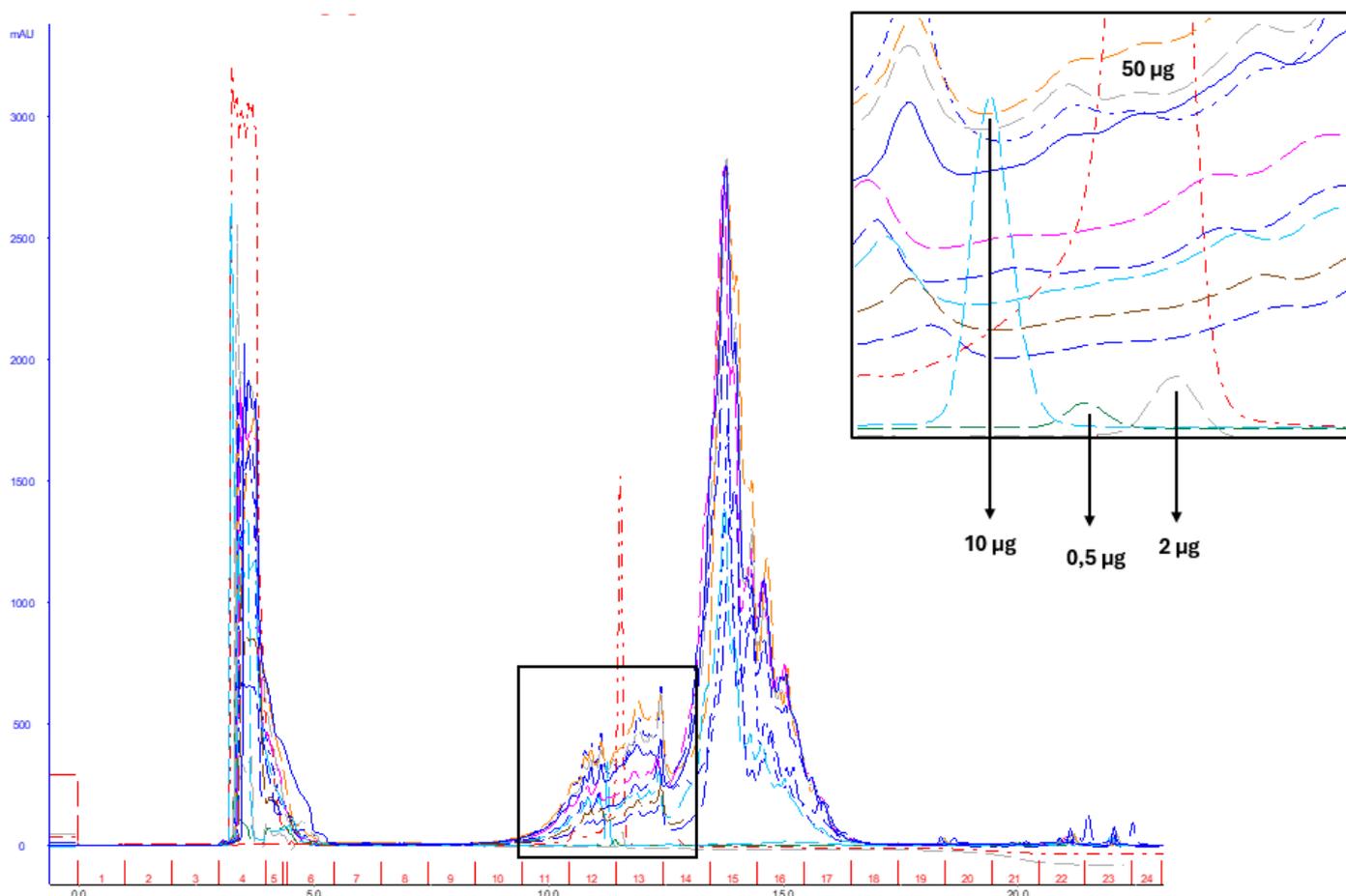
**Figura 1** – Cromatograma da análise de AR por HPLC para amostras E430.



**Nota:** No detalhe, os padrões de 0,5; 2; 10 e 50 ug de AR (Sigma Aldrich). As amostras referem-se às linhagens 4 e 8, além de plantas WT, em triplicata.

**Fonte:** do autor (2025).

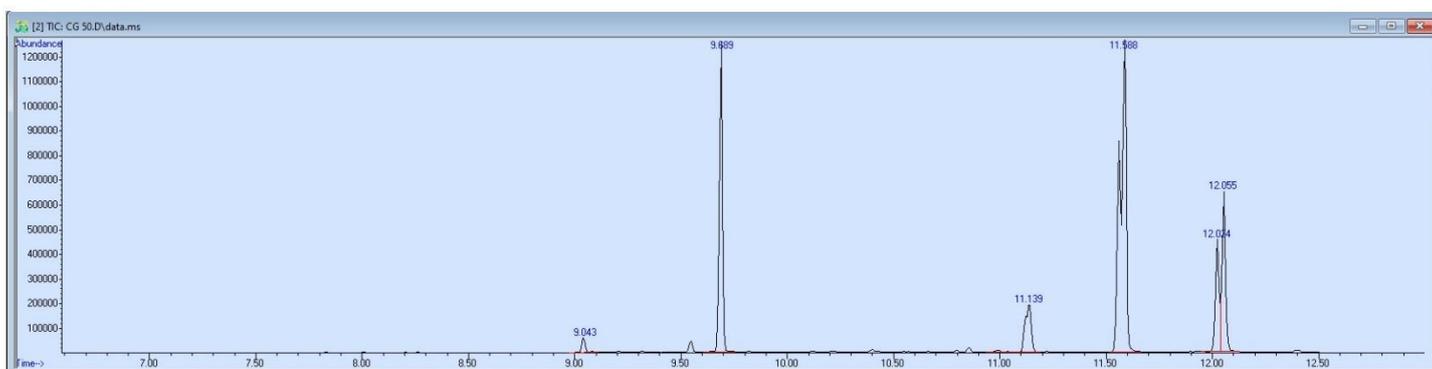
**Figura 2** – Cromatograma da análise de AR por HPLC para amostras E432.



**Nota:** No detalhe, os padrões de 0,5; 2; 10 e 50 ug de AR (Sigma Aldrich). As amostras referem-se às linhagens 1 e 7, além de plantas WT, em triplicata.

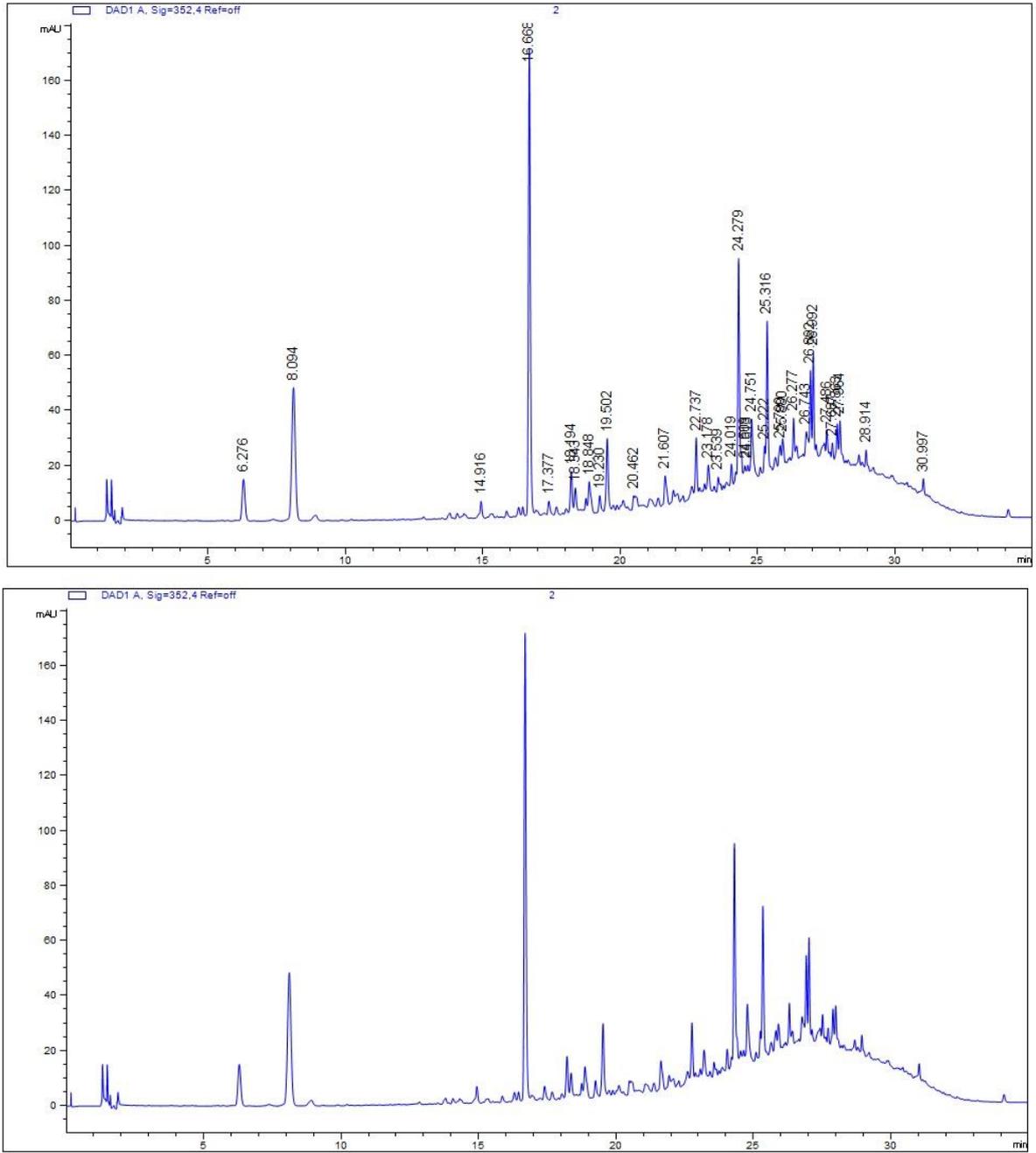
**Fonte:** do autor (2025).

**Figura 3** – Cromatograma CG-MS para tioacidólise.



**Fonte:** do autor (2025).

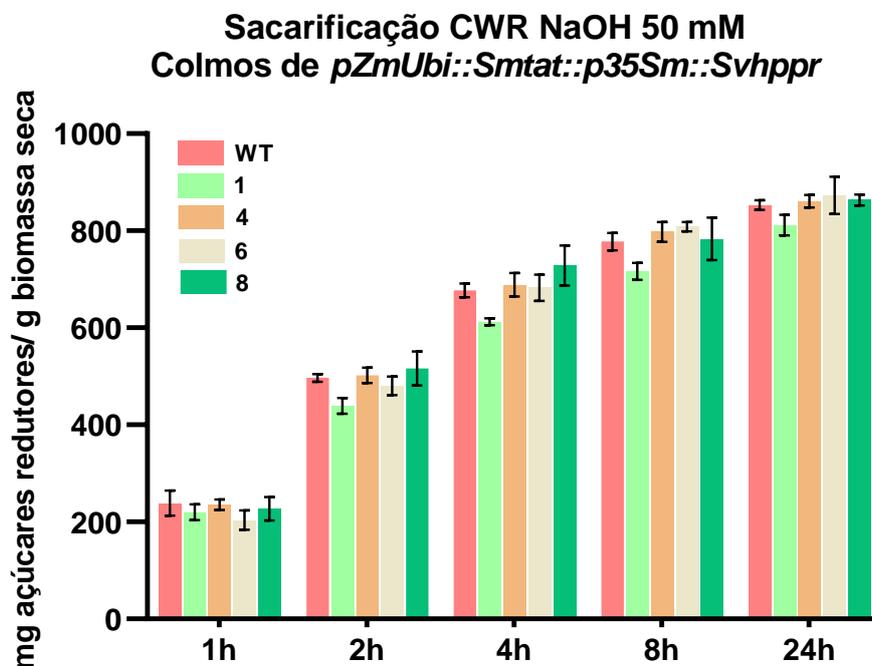
**Figura 4** – Cromatograma para LC-MS para tioacidólise.



**Fonte:** do autor (2025).

## Anexo IV – Sacarificação de CWR com pré-tratamento de E430

**Figura 1** – Sacarificação após pré-tratamento alcalino com NaOH 50 mM de linhagens E430.

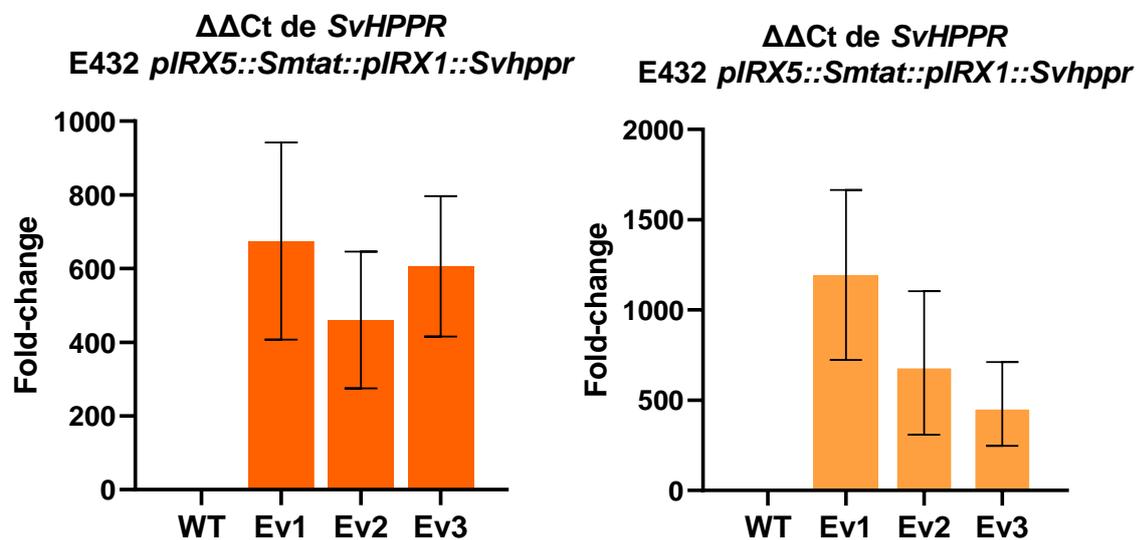


**Nota:** Cada linhagem é representada por pools, sendo pools A, B e C, totalizando um n = 3. Valores acima das barras indicando porcentagem de liberação de açúcar em comparação com WT para cada tempo de hidrólise. Barras de erro se referem ao erro padrão a partir da média (SEM). Asteriscos indicando significância estatística, sendo \*, P < 0.05; \*\*, P < 0.01; \*\*\*, P < 0.001; \*\*\*\*, P < 0.0001.

**Fonte:** do autor (2024).

## Anexo V – Expressão gênica de E432 ao longo do crescimento

**Figura 1** – Expressão gênica de folhas de eventos E432 para Fase I (15 dias) e Fase II (30 dias).



**Nota:** Cada linhagem é representada por pools, sendo pools A, B e C, totalizando um  $n = 3$ . Média dos valores calculados de  $\Delta\Delta Ct$  de cada um dos pools, para cada linhagem. Para WT considerou-se como zero. Barras de erro se referem ao erro padrão a partir da média (SEM).

**Fonte:** do autor (2024).